ARTICLES

Effectiveness of Highly Active Antiretroviral Therapy in HIV-Positive Children: Evaluation at 12 Months in a Routine Program in Cambodia
Bart Janssens, Brian Raleigh, Seithaboth Soeung, Kazumi Akao, Vantha Te, Jitendra Gupta, Mean Chhy Vun, Nathan Ford, Janin Nouhin, and Eric Nerrienet
Pediatrics 2007; 120: e1134-e1140.

Children of Mothers With HIV/AIDS: Unmet Needs for Mental Health Services
Laurie J. Bauman, Ellen J. Silver, Barbara H. Draidin, and Jan Hudis
Pediatrics 2007; 120: e1141-e1147.

Out-of-Pocket Costs of Childhood Immunizations: A Comparison by Type of Insurance Plan
Noëlle-Angélique M. Molinari, Maureen Kolasa, Mark L. Messonnier, and Richard A. Schieber
Pediatrics 2007; 120: e1148-e1156.

Obesity Is a Common Comorbidity in Children With Congenital and Acquired Heart Disease
Nelangi M. Pinto, Bradley S. Marino, Gil Wernovsky, Sarah D. de Ferranti, Amy Z. Walsh, Meena Laronde, Kristen Hyland, Stanley O. Dunn, Jr, and Meryl S. Cohen
Pediatrics 2007; 120: e1157-e1164.

Effect of Vaccine Shortages on Timeliness of Pneumococcal Conjugate Vaccination: Results From the 2001–2005 National Immunization Survey
Philip J. Smith, J. Pekka Nuorti, James A. Singleton, Zhen Zhao, and Kirk M. Wolter
Pediatrics 2007; 120: e1165-e1173.

Does Current Asthma Control Predict Future Health Care Use Among Black Preschool-aged Inner-City Children?
Hemant P. Sharma, Elizabeth C. Matsui, Peyton A. Eggleston, Nadia N. Hansel, Jean Curtin-Brosnan, and Gregory B. Diette
Pediatrics 2007; 120: e1174-e1181.
A Cross-National Comparison of Racial and Ethnic Disparities in Low Birth Weight in the United States and England
Julien O. Teitler, Nancy E. Reichman, Lenna Nepomnyaschy, and Melissa Martinson
Pediatrics 2007; 120: e1182-e1189.

Pertussis Booster Vaccination in HIV-Infected Children Receiving Highly Active Antiretroviral Therapy
Pediatrics 2007; 120: e1190-e1202.

Adherence to Nasal Positive Airway Pressure Therapy Among School-aged Children and Adolescents With Obstructive Sleep Apnea Syndrome
Elizabeth C. Uong, Mary Epperson, Sharon A. Bathon, and Donna B. Jeffe
Pediatrics 2007; 120: e1203-e1211.

Cerebrospinal Fluid Xanthochromia in Newborns Is Related to Maternal Labor Before Delivery
Lise E. Nigrovic, Michelle Trivedi, Jonathan A. Edlow, and Mark I. Neuman
Pediatrics 2007; 120: e1212-e1216.

Is There Differential Retention of Children With Special Health Care Needs in the State Children's Health Insurance Program?
Tamarie Macon, Jane E. Miller, Dorothy Gaboda, Theresa Simpson, and Joel C. Cantor
Pediatrics 2007; 120: e1217-e1224.

Impact of Size at Birth on the Microvasculature: The Avon Longitudinal Study of Parents and Children
Robyn J. Tapp, Cathy Williams, Nicholas Witt, Nish Chaturvedi, Richard Evans, Simon A. McG Thom, Alun D. Hughes, and Andrew Ness

Hearing Loss, Quality of Life, and Academic Problems in Long-term Neuroblastoma Survivors: A Report From the Children's Oncology Group
James G. Gurney, Jean M. Tersak, Kirsten K. Ness, Wendy Landier, Katherine K. Matthy, and Mary Lou Schmidt
Pediatrics 2007; 120: e1229-e1236.

Size at Birth and Motor Activity During Stress in Children Aged 7 to 9 Years
Wolff Schlotz, Alexander Jones, Naomi M.M. Phillips, Keith M. Godfrey, and David I.W. Phillips
Pediatrics 2007; 120: e1237-e1244.
Impact of Singular Excessive Computer Game and Television Exposure on Sleep Patterns and Memory Performance of School-aged Children
Markus Dworak, Thomas Schierl, Thomas Bruns, and Heiko Klaus Strüder

Associations Between Content Types of Early Media Exposure and Subsequent Attentional Problems
Frederick J. Zimmerman and Dimitri A. Christakis

Violent Television Viewing During Preschool Is Associated With Antisocial Behavior During School Age
Dimitri A. Christakis and Frederick J. Zimmerman
Pediatrics 2007; 120: 993-999.

Healing of Nonhymenal Genital Injuries in Prepubertal and Adolescent Girls: A Descriptive Study
John McCann, Sheridan Miyamoto, Cathy Boyle, and Kristen Rogers
Pediatrics 2007; 120: 1000-1011.

Twelve-Month Neurofunctional Assessment and Cognitive Performance at 36 Months of Age in Extremely Low Birth Weight Infants
Maria Lorella Giannì, Odoardo Picciolini, Chiara Vegni, Laura Gardon, Monica Fumagalli, and Fabio Mosca
Pediatrics 2007; 120: 1012-1019.

Shorter Sleep Duration Is Associated With Increased Risk for Being Overweight at Ages 9 to 12 Years
Julie C. Lumeng, Deepak Somashekar, Danielle Appugliese, Niko Kaciroti, Robert F. Corwyn, and Robert H. Bradley

Infant-Feeding Methods and Childhood Sleep-Disordered Breathing
Hawley Evelyn Montgomery-Downs, Valerie McLaughlin Crabtree, Oscar Sans Capdevila, and David Gozal

Pediatricians’ Perspectives Regarding Community Child Health: Training, Involvement, and Expectations According to Age
Cynthia S. Minkovitz, Karen G. O’Connor, Holly Grason, Judith S. Palfrey, Anita Chandra, and Thomas F. Tonniges

Health Status and Health-Related Quality of Life Preference-Based Outcomes of Children Who Are Aged 7 to 9 Years and Have Bilateral Permanent Childhood Hearing Impairment
Stavros Petrou, Donna McCann, Catherine M. Law, Peter M. Watkin, Sarah Worsfold, and Colin R. Kennedy
Pediatrics 2007; 120: 1044-1052.
Association Between Iron-Deficiency Anemia and Stroke in Young Children
Jonathon L. Maguire, Gabrielle deVeber, and Patricia C. Parkin
Pediatrics 2007; 120: 1053-1057.

Impact of Computerized Prescriber Order Entry on the Incidence of Adverse Drug Events in Pediatric Inpatients
Mark T. Holdsworth, Richard E. Fichtl, Dennis W. Raisch, Adrianne Hewryk, Maryam Behta, Elena Mendez-Rico, Cindy L. Wong, Jennifer Cohen, Susan Bostwick, and Bruce M. Greenwald
Pediatrics 2007; 120: 1058-1066.

Ontogeny of Bilirubin-Binding Capacity and the Effect of Clinical Status in Premature Infants Born at Less Than 1300 Grams
George Jesse Bender, William James Cashore, and William Oh

Epidemiology of Oronasal Hemorrhage in the First 2 Years of Life: Implications for Child Protection
Neil McIntosh, Jacqueline Y.Q. Mok, and Adrian Margerison

STATE-OF-THE-ART REVIEW ARTICLE

Effects of Corticosteroid on Henoch-Schönlein Purpura: A Systematic Review
Pamela F. Weiss, James A. Feinstein, Xianqun Luan, Jon M. Burnham, and Chris Feudtner
Pediatrics 2007; 120: 1079-1087.

REVIEW ARTICLES

Keith K. Lau, Jayson M. Stoffman, Suzane Williams, Patricia McCusker, Leonardo Brandao, Sanjay Patel, Anthony K.C. Chan for the Canadian Pediatric Thrombosis and Hemostasis Network
Pediatrics 2007; 120: e1278-e1284.

Inhaled Nitric Oxide for Preterm Infants: A Systematic Review
Keith J. Barrington and Neil N. Finer
SPECIAL ARTICLES

Interpreting and Managing Blood Lead Levels of Less Than 10 µg/dL in Children and Reducing Childhood Exposure to Lead: Recommendations of the Centers for Disease Control and Prevention Advisory Committee on Childhood Lead Poisoning Prevention
Helen J. Binns, Carla Campbell, Mary Jean Brown for the Advisory Committee on Childhood Lead Poisoning Prevention
Pediatrics 2007; 120: e1285-e1298.

Guidelines for Adolescent Depression in Primary Care (GLAD-PC): I. Identification, Assessment, and Initial Management
Rachel A. Zuckerbrot, Amy H. Cheung, Peter S. Jensen, Ruth E.K. Stein, Danielle Laraque and the GLAD-PC Steering Group
Pediatrics 2007; 120: e1299-e1312.

Guidelines for Adolescent Depression in Primary Care (GLAD-PC): II. Treatment and Ongoing Management
Amy H. Cheung, Rachel A. Zuckerbrot, Peter S. Jensen, Kareem Ghalib, Danielle Laraque, Ruth E.K. Stein and the GLAD-PC Steering Group
Pediatrics 2007; 120: e1313-e1326.

Newborn Screening for Pompe Disease: Synthesis of the Evidence and Development of Screening Recommendations
Alex R. Kemper, Wuh-Liang Hwu, Michele Lloyd-Puryear, and Priya S. Kishnani
Pediatrics 2007; 120: e1327-e1334.

Treating Attention-Deficit/Hyperactivity Disorder With a Stimulant Transdermal Patch: The Clinical Art
L. Eugene Arnold, Ronald L. Lindsay, Frank A. López, Sharon E. Jacob, Joseph Biederman, Robert L. Findling, and Yaser Ramadan
Pediatrics 2007; 120: 1100-1106.

COMMENTARIES

Continuity, the Medical Home, and Retail-Based Clinics
Stephen Berman
Pediatrics 2007; 120: 1123-1125.

Cooling for Neonatal Hypoxic Ischemic Encephalopathy: Do We Have the Answer?
Haresh Kirpalani, John Barks, Kristian Thorlund, and Gordon Guyatt
Pediatrics 2007; 120: 1126-1130.
LETTERS TO THE EDITOR

Beyond Munchausen Syndrome by Proxy
Marc D. Feldman, Michael J. Light, Louisa J. Lasher, and Mary S. Sheridan

Beyond Munchausen Syndrome by Proxy: In Reply
John Stirling, Jr and Carole Jenny
Pediatrics 2007; 120: 1218.

Unconventional Solution: A Marketing Approach to Solving Health Literacy Disparities
Marie Leiner and Gilbert Handal
Pediatrics 2007; 120: 1218-1219.

Unconventional Solution: A Marketing Approach to Solving Health Literacy Disparities: In Reply
Lee M. Sanders, James D. Wilkinson, and Valerie Thompson

Parental Attitude to Lumbar Puncture for Children With Fever and Seizure
Ayman Habiba
Pediatrics 2007; 120: 1220.

Parental Attitude to Lumbar Puncture for Children With Fever and Seizure: In Reply
Allison Golnik
Pediatrics 2007; 120: 1221.

Possible Sequelae of Sustained Lung Inflation in Resuscitation of Preterm Infants
Eleanor J. Molloy
Pediatrics 2007; 120: 1221-1222.

Possible Sequelae of Sustained Lung Inflation in Resuscitation of Preterm Infants: In Reply
Arjan B. te Pas and Frans J. Walther
Pediatrics 2007; 120: 1222-1223.

Meta-Analysis Combining 2 Previously Reported Trials on Respiratory Distress Syndrome in Neonates
Philip T. Lavin

Meta-Analysis Combining 2 Previously Reported Trials on Respiratory Distress Syndrome in Neonates
Stephen Senn, Marco Costantini, and Giada Rizzi
Pediatrics 2007; 120: 1224-1225.
Meta-Analysis Combining 2 Previously Reported Trials on Respiratory Distress Syndrome in Neonates: In Reply
Fernando Moya, Sunil Sinha, and Ralph B. D'Agostino, Sr
Pediatrics 2007; 120: 1225-1226.

Relevance of Motor Skill Problems in Victims of Bullying
Susanne Bejerot and Mats Humble
Pediatrics 2007; 120: 1226-1227.

Relevance of Motor Skill Problems in Victims of Bullying: In Reply
Pernille Due, Anette Andersen, Juan Merlo, Ebba Holme Hansen, and Bjørn E. Holstein
Pediatrics 2007; 120: 1227-1228.

EXPERIENCE & REASON

Potential Misdiagnosis of 3-Methylcrotonyl-Coenzyme A Carboxylase Deficiency Associated With Absent or Trace Urinary 3-Methylcrotonylglycine
Lynne A. Wolfe, David N. Finegold, Jerry Vockley, Nicole Walters, Celine Chambaz, Terttu Suormala, Hans Georg Koch, Dietrich Matern, Bruce A. Barshop, Lorna J. Cropcho, Matthias R. Baumgartner, and K. Michael Gibson
Pediatrics 2007; 120: e1335-e1340.

Hematopoietic Stem Cell Transplantation Corrects the Immunologic Abnormalities Associated With Immunodeficiency–Centromeric Instability–Facial Dysmorphism Syndrome
Andrew R. Gennery, Mary A. Slatter, Robbert G. Bredius, Melanie M. Hagleitner, Corry Weemaes, Andrew J. Cant, and Arjan C. Lankester
Pediatrics 2007; 120: e1341-e1344.

Severe Varicella Caused by Varicella-Vaccine Strain in a Child With Significant T-Cell Dysfunction
Patrick Jean-Philippe, Abigail Freedman, Mary Wu Chang, Sharon P. Steinberg, Anne A. Gershon, Philip S. LaRussa, and William Borkowsky
Pediatrics 2007; 120: e1345-e1349.

Neonatal Hyperparathyroidism and Pamidronate Therapy in an Extremely Premature Infant
Lisa Fox, Joel Sadowsky, Kevin P. Pringle, Alexa Kidd, Jean Murdoch, David E.C. Cole, and Esko Wiltshire
Pediatrics 2007; 120: e1350-e1354.

Atypical Tetanus in a Completely Immunized 14-Year-Old Boy
Kai König, Hannelore Ringe, Brigitte G. Dorner, Alexander Diers, Birgit Uhlenberg, Dominik Müller, Verena Varnholt, and Gerhard Gaedicke
Pediatrics 2007; 120: e1355-e1358.
Effectiveness of Highly Active Antiretroviral Therapy in HIV-Positive Children: Evaluation at 12 Months in a Routine Program in Cambodia

Bart Janssens, MD*, Brian Raleigh, MBBS, DTM&H, DipObs*, Seithaboth Soeung, MDb, Kazumi Akaoa, Vantha Te, MDc, Jitendra Gupta, MDb, Mean Chhy Vun, MDb, Nathan Ford, BSc*, Janin Nouhin*, Eric Nerrienet, PhDf

*Me´decins Sans Frontie`res, Phnom Penh, Cambodia; bDepartment of Infectious Diseases, Angkor Hospital for Children, Siem Reap, Cambodia; cDepartment of Pediatrics, Takeo Referral Hospital and dNational Center of HIV/AIDS and Dermatology and STDs, Ministry of Health, Phnom Penh, Cambodia; eMe´decins Sans Frontie`res, Bangkok/HIV/Hepatitis Laboratory, Thailand; fInstitut Pasteur, Cambodia, Phnom Penh, Cambodia

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Increasing access to highly active antiretroviral therapy to reach all those in need in developing countries (scale up) is slowly expanding to HIV-positive children, but documented experience remains limited. We aimed to describe the clinical, immunologic, and virologic outcomes of pediatric patients with >12 months of highly active antiretroviral therapy in 2 routine programs in Cambodia.

METHODS. Between June 2003 and March 2005, 212 children who were younger than 13 years started highly active antiretroviral therapy. Most patients started a standard first-line regimen of lamivudine, stavudine, and nevirapine, using split adult fixed-dosage combinations. CD4 percentage and body weight were monitored routinely. A cross-sectional virologic analysis was conducted in January 2006; genotype resistance testing was performed for patients with a detectable viral load.

RESULTS. Mean age of the subjects was 6 years. Median CD4 percentage at baseline was 6. Survival was 92% at 12 months and 91% at 24 months; 13 patients died, and 4 were lost to follow-up. A total of 81% of all patients had an undetectable viral load. Among the patients with a detectable viral load, most mutations were associated with resistance to lamivudine and non–nucleoside reverse-transcriptase inhibitor drugs. Five patients had developed extensive antiretroviral resistance. Being an orphan was found to be a predictor of virologic failure.

CONCLUSIONS. This study provides additional evidence of the effectiveness of integrating HIV/AIDS care with highly active antiretroviral therapy for children in a routine setting, with good virologic suppression and immunologic recovery achieved by using split adult fixed-dosage combinations. Viral load monitoring and HIV genotyping are valuable tools for the clinical follow-up of the patients. Orphans should receive careful follow-up and extra support.
Access to antiretroviral treatment for patients with AIDS in resource-limited settings has increased dramatically in the past few years, with >1.6 million currently estimated to be receiving treatment. Despite these efforts, however, <5% of the 800,000 children who are in need of antiretroviral therapy are receiving it.1 Of the 39.5 million people who were estimated to be living with HIV/AIDS in 2006, 2.3 million were children who were younger than 15; more than half a million children died of HIV/AIDS that year.2 In the West, good outcomes are achieved for children who are on highly active antiretroviral therapy (HAART),3–7 and studies indicate that similar outcomes can be achieved in resource-limited settings.8–11 However, documented experience of treating large numbers of children remains limited.

Cambodia is 1 of the poorest countries in Asia, ranking 130th on the human development index with an annual gross domestic product per capita of $2078.12 Generally the prevention response has been good in Cambodia, with HIV prevalence falling from 3% in 1997 to 1.9% in 2003. However, national efforts to prevent mother-to-child transmission only began in 2003.

Antiretroviral therapy has been provided in Cambodia since 2001. By December 2005, government and nongovernmental actors supported ~12,000 patients with HAART, which allowed the country to reach its World Health Organization (WHO) 3 by 5 target, which aims to start HAART for 3 million patients with AIDS worldwide by the end of 2005. Children have benefited from treatment since it has been available, with Médecins Sans Frontières (MSF) and the Cambodian Ministry of Health providing HAART to children since 2002. This was given additional priority in 2004, when the Ministry of Health released a policy package of treatment guidelines, training, and drug supply devoted to the management of pediatric HIV. This article provides outcomes of providing HAART for 212 children within a cohort of >800 HIV-positive children who were under care in 2 district hospitals in Cambodia.

Methods

Study Sites and Patients

Outcome data were pooled from 2 district hospitals: Angkor Hospital for Children in Siem Reap province (population 700,000) and Takeo district hospital in Takeo province (population 800,000). Angkor Hospital for Children is a charity-run 50-bed pediatric hospital that started to provide care for HIV-positive children in collaboration with MSF in February 2003. The Takeo district hospital is a 164-bed public referral hospital that started providing HIV care for children in February 2004, also with support from MSF. Both hospitals attract children from across the province as well as from neighboring provinces. Services in both sites are comparable and are provided primarily by Cambodian medical staff. The support of MSF was limited to technical assistance, the supply of most of the medicines, and the organization of the laboratory tests.

By March 2006, a total of 805 HIV-positive children (478 in Siem Reap and 327 in Takeo) had been enrolled in these 2 sites, 428 of whom had started HAART. Of the remainder, 82 (10%) patients were eligible for HAART but had died or were lost to follow-up before initiation and 295 (37%) did not meet eligibility criteria for HAART (see “Antiretroviral Therapy” below). All children at both sites who were aged ≤13 years and on HAART for >12 months were included in a cross-sectional evaluation of the virologic status, with an intention-to-treat analysis of the virologic efficacy performed on all children who started HAART between June 2003 and March 2005.

Antiretroviral Therapy

Criteria to commence HAART followed national guidelines: HIV-positive status (serology or reverse transcriptase polymerase chain reaction for children younger than 18 months) and CD4 count (<20% for children younger than 5 years or <200 cells per mL for children older than 5 years). The majority of the patients started on a standard first-line regimen of stavudine, lamivudine, and nevirapine, as recommended by the WHO.13 Zidovudine and efavirenz were used as alternatives in case of intolerance or interaction with other drugs. Adult generic fixed-dosage combinations (FDCs) were used for the children who were on the first-line regimen: before December 2003, GPOvir (GPO, Bangkok Thailand) was used; patients were then switched to Triviro (Ranbaxy, Dewas, India) on prequalification of this treatment.14 Tablets were cut in half to obtain the most appropriate dosages, and for some body weights, nevirapine syrup was added to achieve the correct dosage, according to a standardized drug-dosage table.15 For the patients in whom treatment failure was diagnosed, an individual evaluation was made to decide on a switch to a second-line HAART regimen. All patients received at least 3 HAART-preparation counseling sessions before commencing the treatment. For younger children, counseling was given in the presence of the parents or the caregivers; for older children, a part of the sessions were given to the patients alone. Disclosure of HIV status was not a condition to start HAART, but the counselors developed an individual plan for each patient to follow a process of disclosure. Adherence to HAART was not systematically recorded, but several steps were taken to provide support for adherence. After commencement of HAART, adherence support was provided by the counselors at every visit to the clinic. Home visits were conducted for patients who arrived late for appointments or were suspected of needing more social support (although this was not always feasible for all patients who
lived in another province). For patients who could not afford the travel to the clinics, financial support for transportation cost was provided.

Clinical, Immunologic, and Virologic Evaluation
Body weight was measured at every visit. CD4 cell count was measured at baseline and every 6 months after HAART initiation by flow cytometry (Facscount; Becton Dickinson, San Jose, CA). For children who were younger than 5 years, CD4% was used as the standard follow-up indicator; for the older children, both CD4% and the absolute CD4 cell count were used as recommended by the WHO.13 Hemoglobin and alanine aminotransferases were measured at week 2; months 1, 2, 3, and 6; and every 6 months thereafter to monitor for toxicity.

Between February 2006 and June 2006, viral load measurements were performed together with routine CD4 tests (flow cytometry, Facscount) for all patients who were on HAART for ≥12 months. HIV-1 RNA viral load was quantified by real-time reverse transcriptase polymerase chain reaction as developed by the French “Agence Nationale de Recherche sur le SIDA et les Hépatites Virales.” For patients with a detectable viral load (ie, >400 copies per mL), a genotype resistance test was done. For the genotype resistance study, viral RNA was extracted (QIampViral RNA minikit; Qiagen, Hamburg, Germany) and amplified for reverse transcriptase genes, sequenced automatically (CEQ DTCs Quick start kit; Beckman Coulter, Fullerton, CA), and corrected manually using CEQ 8000 software (Beckman Coulter). Nucleotide sequences were compared against known reference strains of group M of the HIV-1 gene bank (www.hiv-web.lanl.gov). Genotype resistance interpretation was performed according to algorithms developed by the “Agence Nationale de Recherche sur le SIDA et les Hépatites Virales” (www.hivfrenchresistance.org) and the International AIDS Society (www.iasusa.org). All viral load and genotype resistance tests were performed at the Pasteur Institute (Phnom Penh, Cambodia).

This cohort evaluation received ethical approval of the Angkor Hospital for Children ethics committee; all patients were informed about the objectives of the laboratory tests during the routine counseling sessions. The results of all of the tests were made available to the clinicians as soon as possible to guarantee optimal benefits for the patients.

Statistical Analysis
All data regarding medical background and follow-up of patients were collected on FUCHIA 1.5 monitoring software (Epicenter, Paris, France). The z score for weight for height to evaluate growth was used from the WHO (Geneva, Switzerland) reference curves. The probability of survival and remaining in care was calculated with Kaplan-Meier with all patients analyzed on an intention-to-treat basis. Patients who were alive and in care or transferred out on March 31, 2006, were censored on the date of their last visit before this date. Patients with >3 months’ delay with respect to the scheduled follow-up visit were regarded as lost to follow-up. A survival analysis to estimate the probability of patients’ remaining in care was done with events (uncensored) counted as patients who died or who were lost to follow-up. Confidence intervals (CIs) around the proportions of patients whose treatment was successful or failing were calculated with Mantel-Haenszel test with 95% limits. Logistic regression analysis was used to identify factors that were associated with virologic failure (viral load > 400 copies per mL). For all listed baseline factors, both univariate and multivariate analyses were calculated. All statistical analyses were performed by using Epi Info 2002 software (Centers for Disease Control and Prevention, Atlanta, GA).

RESULTS
Characteristics of the Study Population
A total of 212 patients had started HAART between June 2003 and March 2005. Baseline characteristics are presented in Table 1. Sixteen patients had started HAART in 2003, 131 in 2004, and 65 in the first quarter of 2005. Median age at start of HAART was 6 years (interquartile range [IQR]: 4.2–7.9), and only 7 (3%) children were younger than 18 months. The primary caregivers were biological parents (64%); grandparents (30%); or relatives, neighbors, or orphans (6%). A total of 76 (36%) of the children were orphans. Only 8 (10.5%) of the 76 orphans were cared for in a structured orphanage. Median CD4% before starting HAART in all patients was 6% (IQR: 2.6–13.0), and the median absolute CD4 count for the children aged ≥5 years (n = 134) was 100 cells per mL (IQR: 22–273). In January 2006, 193 of the patients were alive and could be included in the cross-sectional virologic evaluation.

Treatment Outcomes
Of 212 children who had started HAART before March 2005, 13 had died, 4 were lost to follow-up, and 2 were transferred to another treatment site. Tuberculosis was responsible for 8 deaths; the remainder were attributed to severe sepsis (2 deaths), pneumonia, bacterial meningitis, and wasting syndrome. According to the survival analysis, 92% of the children were alive and in care at 12 months and 91% at 24 months of treatment (Fig 1). Median time on HAART of the patients who were alive was 16.8 months (IQR: 13.9–21.2). Thirty-six (17%) patients were on HAART for ≥24 months. Mean weight-for-height z score increased by 0.81 to −0.78 (SD: ±1.17) at 12 months of HAART. HAART was generally well tolerated, with only 7 patients having to switch to
TABLE 1 Baseline Characteristics of Cambodian Children in Takeo and Siem Reap Who Started HAART Before March 2005

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
</tr>
<tr>
<td>All patients</td>
<td>212</td>
</tr>
<tr>
<td>Female, n (%).</td>
<td>94 (44.4)</td>
</tr>
<tr>
<td>Age, median (IQR), y</td>
<td>6 (4–7.9)</td>
</tr>
<tr>
<td>Clinicoinmunology</td>
<td></td>
</tr>
<tr>
<td>CDC stage at HAART initiation, n (%).</td>
<td></td>
</tr>
<tr>
<td>Stage A (asymptomatic), n (%)</td>
<td>20 (9.5)</td>
</tr>
<tr>
<td>Stage A (mild), n (%)</td>
<td>55 (26)</td>
</tr>
<tr>
<td>Stage B (moderate), n (%)</td>
<td>93 (43.8)</td>
</tr>
<tr>
<td>Stage C (severe), n (%)</td>
<td>44 (20.7)</td>
</tr>
<tr>
<td>Weight-for-height z score, mean ± SD</td>
<td>−1.59 ± 1.08</td>
</tr>
<tr>
<td>CD4 cells (all patients), median (IQR), %</td>
<td>6 (2.6–13)</td>
</tr>
<tr>
<td>CD4 cells &lt; 5% (all patients), %</td>
<td>84 (40)</td>
</tr>
<tr>
<td>CD4 count (patients ≥ 5 y), median (IQR), cells per mL (n = 134)</td>
<td>100 (22–273)</td>
</tr>
<tr>
<td>Initial antiretroviral treatment</td>
<td></td>
</tr>
<tr>
<td>D4T/3TC/NVP, n (%)</td>
<td>146 (68.9)</td>
</tr>
<tr>
<td>D4T/3TC/EFV, n (%)</td>
<td>40 (18.8)</td>
</tr>
<tr>
<td>AZT/3TC/NVP, n (%)</td>
<td>26 (12.3)</td>
</tr>
<tr>
<td>Time on HAART, median (IQR), mo</td>
<td>16.8 (13.9–21.2)</td>
</tr>
<tr>
<td>Social, n (%)</td>
<td></td>
</tr>
<tr>
<td>Geographic origin</td>
<td></td>
</tr>
<tr>
<td>From the same province as the clinic</td>
<td>90 (42.4)</td>
</tr>
<tr>
<td>From another province</td>
<td>122 (67.6)</td>
</tr>
<tr>
<td>Caregiver</td>
<td></td>
</tr>
<tr>
<td>At least 1 of the parents</td>
<td>136 (64)</td>
</tr>
<tr>
<td>Orphans</td>
<td>76 (36)</td>
</tr>
<tr>
<td>Grandparents as caregiver</td>
<td>63 (30)</td>
</tr>
<tr>
<td>Other caregiver</td>
<td>13 (6)</td>
</tr>
<tr>
<td>Outcomes, n (%)</td>
<td></td>
</tr>
<tr>
<td>Active in treatment</td>
<td>193 (91)</td>
</tr>
<tr>
<td>Died</td>
<td>13 (6.1)</td>
</tr>
<tr>
<td>Lost to follow-up</td>
<td>4 (2)</td>
</tr>
<tr>
<td>Transferred to another treatment site</td>
<td>2 (0.9)</td>
</tr>
<tr>
<td>Mortality rate, per person-year</td>
<td>4</td>
</tr>
</tbody>
</table>

CDC indicates Centers for Disease Control and Prevention; D4T, stavudine; 3TC, lamivudine; NVP, nevirapine; EFV, efavirenz; AZT, zidovudine.

TAB21021

an alternative regimen for reasons of intolerability for nevirapine (2), zidovudine (2), and stavudine (3).

Median CD4% gain at 12 months was 17.0% (IQR: 16.3%–30.7%); among the children who were older than 5 years at that time (n = 164), median CD4 cell count gain from baseline was 490 cells per mL. Viral load was measured for all 193 patients who were alive and on HAART for >12 months. A total of 156 (81% [95% CI: 74%–85%]) of the samples showed an undetectable viral load (<400 copies per mL), 7 (3.7% [95% CI: 1.6%–7.5%]) of the patients had a viral load between 400 and 1000 copies per mL, and 30 (15.5% [95% CI: 11.0%–21.4%]) of the patients had a viral load of >1000 copies per mL.

In an intention-to-treat analysis, 156 of 212 (73.6% [95% CI: 67.0%–79.0%]) of all patients who started HAART before March 2005 had an undetectable viral load. Of these, 36 had been on HAART for 24 months, and among this subgroup, 24 (66.6% [95% CI: 50.0%–80.0%]) had an undetectable viral load. The proportion with an undetectable viral load in the intention-to-treat analysis for this group was 24 of 38 (63.0% [95% CI: 47.0%–76.6%]). Of the 30 children with a viral load of >1000 copies, 16 (53%) were in immunologic failure at the moment of the viral load according to the criteria of the guidelines from the National Institutes of Health.16

In the analysis of baseline characteristics as predictors of virologic failure (Table 2), being an orphan was found to be a statistically significant predictor of virologic failure (P = .001). For all 37 samples with a viral load of >400 copies per mL, genotype resistance testing was performed. Genotyping could not be amplified in 1 case. In 34 samples, at least 1 mutation was found. Of the 2 samples in which none was detected, viral loads were 16 637 and 689 copies per mL. Mutations in 31 of 34 patients showed a nucleoside reverse-transcriptase inhibitor (NRTI) mutation, the most common being M184 (n = 26), D67 (n = 7), T215 (n = 7), and T69 and L210 (n = 4). Twenty-seven patients had developed resistance to lamivudine, 9 to zidovudine, and 11 to stavudine.

Resistance to abacavir was found in 4 patients, to tenofovir in 3 patients, and to didanosine in 3 patients. Thirty-two of 34 had developed non–nucleoside reverse-transcriptase inhibitor (NNRTI) resistance, 3 of them had no NRTI resistance. Of 31 patients with NRTI resistance, 2 showed no NNRTI resistance.

DISCUSSION
The experience of these 2 pediatric HIV/AIDS care programs shows that antiretroviral therapy is feasible for children at the district level in Cambodia, with very good results obtained after 2 years on treatment. A total of 92% of patients were alive at 12 months, and median CD4% rose from 6% at baseline to 25% at 12 months. A cross-sectional survey of viral load of all patients who were receiving HAART at ≥ 12 months showed that 81% had an undetectable viral load using a threshold of detection of 400 copies per mL as treated and 74% in intention-to-treat analysis.

The proportion of patients who achieved virologic success is equal to or better than all of the clinical trials in Western settings before 2002 as summarized in a review article by Van Rossum et al.17 Among the best results cited in the review, 3 trials are comparable; they report the proportion of patients with a viral load of <400 copies per mL at 48 to 72 weeks of HAART as 69%, 70%, and 87%, respectively, on a population as treated. Only 1 result is available on an intention-to-treat population, and that showed 61% with a viral load of <400 copies per mL. In the comparison with the results of our Cambodian cohort, it has to be acknowledged, however, that in a number of these studies, no patients were included that were naïve to NRTI. The outcomes in this study generally also compare favorably
with the published cohort studies in other resource-limited contexts such as Thailand, Romania, and Ivory Coast.8,10,11

The majority of patients in our cohort were treated with split adult FDC of 2 NRTIs and 1 NNRTI, confirming other findings that good outcomes can be achieved with this treatment strategy.18,19 It should be noted, however, that this practice is a suboptimal interim strategy pending prequalification of pediatric formulations of FDCs. The majority of the patients were in an advanced stage of HIV infection, similar to many settings where HIV/AIDS care for children has been commenced. We also note that the 2 different settings offer similar results, demonstrating that a well-organized public hospital (Takeo) with targeted support can achieve a quality of care equal to a private, nonprofit setting. In Takeo, all consultations and nursing support was done by Ministry of Health hospital staff, and MSF provide psychosocial support and technical support.

Being orphaned was found to lead to a greater risk for virologic failure, and this is likely because of poorer adherence support, although this cannot be said with certainty given that adherence was not recorded systematically. Most of these orphans are cared for in the family of the grandparents or neighbors. The results of this evaluation should encourage paying a greater attention to adherence support for this group, for example, through more intensive home-based care, specific counseling sessions for grandmothers, or the development of better adapted tools and approaches for adherence support.

We are aware of 7 patients who had previous antiretroviral experience (purchased drugs from private pharmacies or shared with their parents). All but 1 of these developed virologic treatment failure. We do not know for certain whether other children were also an-

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Risk Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td></td>
</tr>
<tr>
<td>≤6</td>
<td>1.00</td>
</tr>
<tr>
<td>&gt;6</td>
<td>1.200 (0.580–2.480)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.00</td>
</tr>
<tr>
<td>Male</td>
<td>1.900 (0.910–4.100)</td>
</tr>
<tr>
<td>Adjusted z score for weight (per unit increase)</td>
<td>0.867 (0.612–1.230)</td>
</tr>
<tr>
<td>CD4 %</td>
<td></td>
</tr>
<tr>
<td>&gt;5</td>
<td>1.00</td>
</tr>
<tr>
<td>≤5</td>
<td>0.685 (0.330–1.400)</td>
</tr>
<tr>
<td>CDC stage</td>
<td></td>
</tr>
<tr>
<td>All other stages</td>
<td>1.00</td>
</tr>
<tr>
<td>Stage C</td>
<td>1.700 (0.560–5.280)</td>
</tr>
<tr>
<td>Geographic origin</td>
<td></td>
</tr>
<tr>
<td>Patients from the same province</td>
<td>1.00</td>
</tr>
<tr>
<td>Patients from outside the province of the clinic</td>
<td>1.230 (0.558–2.730)</td>
</tr>
<tr>
<td>Caregiver</td>
<td></td>
</tr>
<tr>
<td>Children with at least 1 parent</td>
<td>1.00</td>
</tr>
<tr>
<td>Orphan</td>
<td>3.850 (1.830–8.110)</td>
</tr>
<tr>
<td>Regimen at start</td>
<td></td>
</tr>
<tr>
<td>AZT/3TC/NVP</td>
<td>1.00</td>
</tr>
<tr>
<td>D4T/3TC/NVP</td>
<td>1.180 (0.315–4.450)</td>
</tr>
<tr>
<td>D4T/3TC/EFV</td>
<td>1.110 (0.411–2.970)</td>
</tr>
</tbody>
</table>

Multivariate values were only represented when univariate values showed a significant effect of the baseline characteristic.

The majority of patients in our cohort were treated with split adult FDC of 2 NRTIs and 1 NNRTI, confirming other findings that good outcomes can be achieved with this treatment strategy.18,19 It should be noted, however, that this practice is a suboptimal interim strategy pending prequalification of pediatric formulations of FDCs. The majority of the patients were in an advanced stage of HIV infection, similar to many settings where HIV/AIDS care for children has been commenced. We also note that the 2 different settings offer similar results, demonstrating that a well-organized public hospital (Takeo) with targeted support can achieve a quality of care equal to a private, nonprofit setting. In Takeo, all consultations and nursing support was done by Ministry of Health hospital staff, and MSF provide psychosocial support and technical support.

Being orphaned was found to lead to a greater risk for virologic failure, and this is likely because of poorer adherence support, although this cannot be said with certainty given that adherence was not recorded systematically. Most of these orphans are cared for in the family of the grandparents or neighbors. The results of this evaluation should encourage paying a greater attention to adherence support for this group, for example, through more intensive home-based care, specific counseling sessions for grandmothers, or the development of better adapted tools and approaches for adherence support.

We are aware of 7 patients who had previous antiretroviral experience (purchased drugs from private pharmacies or shared with their parents). All but 1 of these developed virologic treatment failure. We do not know for certain whether other children were also an-

FIGURE 1
Survival of children who started HAART in Takeo and Siem Reap.

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Effects of the Baseline Characteristics on the Likelihood That Plasma HIV RNA Would be &gt;400 Copies per mL After ≥12 Months of HAART</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>Risk Ratio (95% CI)</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
</tr>
<tr>
<td>≤6</td>
<td>1.00</td>
</tr>
<tr>
<td>&gt;6</td>
<td>1.200 (0.580–2.480)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.00</td>
</tr>
<tr>
<td>Male</td>
<td>1.900 (0.910–4.100)</td>
</tr>
<tr>
<td>Adjusted z score for weight (per unit increase)</td>
<td>0.867 (0.612–1.230)</td>
</tr>
<tr>
<td>CD4 %</td>
<td></td>
</tr>
<tr>
<td>&gt;5</td>
<td>1.00</td>
</tr>
<tr>
<td>≤5</td>
<td>0.685 (0.330–1.400)</td>
</tr>
<tr>
<td>CDC stage</td>
<td></td>
</tr>
<tr>
<td>All other stages</td>
<td>1.00</td>
</tr>
<tr>
<td>Stage C</td>
<td>1.700 (0.560–5.280)</td>
</tr>
<tr>
<td>Geographic origin</td>
<td></td>
</tr>
<tr>
<td>Patients from the same province</td>
<td>1.00</td>
</tr>
<tr>
<td>Patients from outside the province of the clinic</td>
<td>1.230 (0.558–2.730)</td>
</tr>
<tr>
<td>Caregiver</td>
<td></td>
</tr>
<tr>
<td>Children with at least 1 parent</td>
<td>1.00</td>
</tr>
<tr>
<td>Orphan</td>
<td>3.850 (1.830–8.110)</td>
</tr>
<tr>
<td>Regimen at start</td>
<td></td>
</tr>
<tr>
<td>AZT/3TC/NVP</td>
<td>1.00</td>
</tr>
<tr>
<td>D4T/3TC/NVP</td>
<td>1.180 (0.315–4.450)</td>
</tr>
<tr>
<td>D4T/3TC/EFV</td>
<td>1.110 (0.411–2.970)</td>
</tr>
</tbody>
</table>

Multivariate values were only represented when univariate values showed a significant effect of the baseline characteristic.
tiretroviral experienced before they started follow-up in the clinics, so it is not possible to calculate the treatment success rate among antiretroviral-naïve patients. In the MSF cohorts of adult patients in Cambodia (>5000 patients), ~5% of all new patients are antiretroviral experienced, and a large proportion of pediatric patients have parents on HAART, so it is possible that more pediatric patients are antiretroviral experienced.

Sixteen of 30 patients with a viral load of >1000 copies per mL showed clear immunologic failure at the last CD4 measurement. Previous studies have concluded that CD4 monitoring is more appropriate than virologic monitoring because a decreasing CD4 count is a better predictor of disease progression. Early detection of treatment failure is important to limit the selection of resistance mutations, although it remains unclear at which level of detectable viral load patients should be switched.20 All but 1 patient with a viral load of >1000 copies per mL had developed ≥1 resistance mutation indicating failure to the first-line HAART regimen. Seven patients had developed a profile of extensive resistance, leaving very few options for a HAART regimen that could be effective in achieving viral suppression. Five of these patients were known to be antiretroviral experienced. Five of the 30 patients in whom treatment failure was detected had virus strains that were sensitive to only 1 drug in the suggested standard second-line regimen in the 2006 WHO guidelines for infants and children.13 Three of these patients were known to be antiretroviral experienced, which suggests that the use of a standard second-line regimen will not be a good solution for all patients, and an expanded formulary is needed.

There are several limitations to this study. First, although a small group of patients have been on HAART for at least 24 months, the overall time under HAART is still short. Second, we do not have good information on antiretroviral experience for all of the patients; neither have we attempted to measure adherence in a systematic way. Third, only 7 children who started HAART at younger than 18 months are included in this cohort because of the unavailability of appropriate diagnostic tools to diagnose HIV/AIDS in infants in the first years of the clinics’ activities, and only rarely did children with suspected HIV infection reach the clinics before the age of 1.

Despite these limitations, these findings provide clear support for the feasibility of integrating pediatric care in a HAART clinic at the district level. Viral load is a valuable tool for detecting early treatment failure, and genotyping is useful for choosing the best second-line regimen, particularly for treatment-experienced patients. Orphans are at higher risk for developing treatment failure and need extra attention and support.

ACKNOWLEDGMENTS
Both projects were funded by MSF and Angkor Hospital for Children.

We thank all patients, caregivers, and staff at Angkor Hospital for Children and the Takeo Hospital pediatric ward; the laboratory staff at the Pasteur Institute of Cambodia for the viral loads and the HIV-resistance tests; and Dr Myrto Schaefer and Dr Saphonn Vonthanak for valuable review of the text.

REFERENCES
Children of Mothers With HIV/AIDS: Unmet Needs for Mental Health Services

Laurie J. Bauman, PhD, Ellen J. Silver, PhD, Barbara H. Draimin, DSW, Jan Hudis, MPH

*Department of Pediatrics, Albert Einstein College of Medicine/Children’s Hospital at Montefiore, Bronx, New York; †Family Center, New York, New York

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The objective of this study was to use multiple measures and sources to assess mental health over time in uninfected 8- to 12-year-old children of HIV-positive mothers.

METHODS. We recruited from the New York City Division of AIDS Services and Income Support a consecutive sample of 157 single mothers who were living with an HIV-negative child aged 8 to 12. Families were randomly assigned to receive a custody planning intervention, Project Care, or standard care. Data were collected at baseline and 4 subsequent times 6 months apart. Children completed the Children’s Depression Inventory; 8- to 10-year-olds completed the Terry, and 11-to 12-year-olds completed the Youth Self-Report. Mothers completed the Child Behavior Checklist. Each measure has a validated cutoff score to signify clinically significant symptoms.

RESULTS. All 5 data points were available on 129 (82%) children. During 2 years, every child had a score in the clinical range (12% once, 25% twice, 26% 3 times, 27% 4 times, and 9% all 5 times). Clinically significant symptoms were most likely at baseline when mothers were sickest. Few had clinically significant symptoms based on maternal report only (5%) or child report only (8%). Chronicity of clinically significant symptoms was not related to child age or gender, maternal health or depression, parent-child relationship, or being assigned to Project Care. Although two thirds of the children received mental health services during the study, <25% did at any 1 time, and 28% of children with chronic clinically significant symptoms never received care.

CONCLUSIONS. Children who are affected by AIDS should be routinely screened for psychiatric problems by using multiple measures and sources to avoid underidentification and be carefully monitored long-term.
M ost of the women who are living with HIV/AIDS in the United States are mothers1 whose children rely on them for their care.2 These children face extraordinary challenges to their mental health and adjustment. Their parents may be too ill to attend to their needs; almost half (45%) of mothers have symptoms consistent with a psychiatric disorder, a known risk factor for poor child mental health; 10% need drug or alcohol treatment; and 1 in 4 lacks key elements of social support.1

Changes in the HIV pandemic have important social and psychological implications for children of mothers with HIV/AIDS. The number of at-risk children whose mothers have HIV/AIDS is likely to grow in the next decade. First, the length of survival of people with HIV/AIDS has increased dramatically with the introduction of highly active antiretroviral therapy,3 and children will spend more of their formative years in a home with an HIV-positive mother. Second, 84% of women with HIV are of childbearing age,4 and most acquired the virus through unprotected heterosexual intercourse.3 New pregnancies are likely to increase the number of children who are cared for by infected women. Third, the number of new cases of HIV/AIDS among women is growing; the HIV epidemic is spreading at a faster rate among women than men, especially among younger women.5 Although it is still necessary to plan for the increasing numbers of AIDS orphans,6,7 we must also acknowledge the increasing number of HIV-infected mothers who will be raising uninfected children.

Children of mothers with HIV/AIDS may be at high risk for psychological disturbance. Studies on the mental health of HIV-negative children of HIV-positive mothers have focused on internalizing (eg, depression, anxiety) and externalizing (eg, conduct problems) symptoms and have found that children of HIV-positive mothers tend to exhibit heightened levels of both types of problems.8-18 Studies that have examined attention problems and cognitive and social competence have also found that children of parents with HIV have more problems in these areas.11-13 However, some research has not shown an association between parental HIV/AIDS and problems in child functioning.19-22

Existing literature on the mental health of children whose parents have HIV is limited in several ways. Most studies have relatively small sample sizes and large age ranges, use cross-sectional data, and rely on parent or caregiver report only. The goal of this study was to measure the mental health of children who were aged 8 to 12 and whose mothers had late stage HIV/AIDS using multiple measures, 2 sources of information, and data during a 2-year period. Several studies have reported the cumulative prevalence of child psychiatric diagnoses in general populations of children not affected by parental HIV. Costello et al23 found that 36.7% of children aged 9 to 13 had received at least 1 psychiatric diagnosis by age 16; similar rates were found in studies done in New Zealand24 (38.3%) and New York25 (39%) using different populations and measures. Other children at high risk have been documented to have much higher rates of disorder than the general population of children. For example, estimates of psychiatric disorder in low-income children ranges from 17% to 26%26-28; in homeless children from 32% to 78%,29,30 and for children in foster care from 54% to 80%.31,32 We hypothesized that inner-city children of mothers with HIV/AIDS would have more risk factors for poor mental health and demonstrate higher burden of psychological symptoms.

METHODS

We recruited a consecutive sample of 220 women from the New York City Division of AIDS Services Income Support (DASIS). New York City provides a package of financial services, housing, and home care to people who have HIV or AIDS and meet income criteria (eligibility for Medicaid in New York State) and medical criteria (including diagnosis of AIDS or late-stage HIV disease and need for home care assistance). DASIS estimated that >95% of eligible people with AIDS actually receive services; therefore, our sample of women was reasonably representative of poor mothers with late-stage HIV/AIDS in New York City.

We recruited women who first entered DASIS services between June 1996 and June 1998 to participate in Project Care, a 2-group randomized trial of a custody-planning intervention provided by the staff of The Family Center.31 Data for this article were from the baseline interview, conducted before receiving any Project Care services, and from the 4 follow-up interviews conducted at 6, 12, 18, and 24 months after study entry (5 time points total). To be eligible for Project Care, mothers had to qualify for DASIS services, have an HIV-negative child who was between the ages of 2 and 12 years and in her custody and whose biological father did not reside in the home, be English or Spanish speaking, have access to a telephone, and not be homeless. We established eligibility on all mothers with children aged 2 to 12 and successfully interviewed 93% of these. In addition, in households with multiple uninfected children aged 8 to 12 years, 1 child was randomly selected to be interviewed (n = 157). Human subjects approval was obtained from the institutional review boards at the Albert Einstein College of Medicine and the Medical and Health Research Association of New York City. Informed consent was obtained from the mother, and assent was obtained from the child. Each mother also signed a New York State Department of Health Authorization for the Release of HIV-related Confidential Information to acknowledge that she knew that The Family Center and the researchers were aware of her HIV status and would not disclose it to anyone else. This study uses data from the 157 mother-child dyads. During the 2-year follow-up period, we lost 28 families to death or attrition,
leaving a final sample of 129 (82%). Neither attrition rates nor background characteristics of those who were lost to follow-up differed between the experimental and control groups.

**Measures**

**Child Mental Health**

We used multiple measures and both maternal and child report to assess child mental health. Mothers completed the 118-item Child Behavior Checklist (CBCL)\(^34\) of behavioral problems that provides a total score as well as internalizing and externalizing subscales and 9 symptom subscales. We used the recommended T-score transformations of the raw behavior scores, which adjust for age and gender differences in behavior found in normative samples. A T score of ≥63 on the total score (the highest 10th percentile) or a score above the 97th percentile on any individual subscale indicates clinically meaningful symptoms.

All children completed the Children’s Depression Inventory (CDI).\(^35,36\) This self-report scale has 27 items and is designed for use in school-aged children and adolescents. Cronbach’s α has ranged from .71 to .87 in various studies and has been demonstrated to be reliable.\(^36\) The CDI is the most widely used self-report measure of depression in children. Normative data are available from psychiatric, pediatric, and school-based populations. A CDI score of ≥12 was used as the cut point to indicate clinically significant levels of symptoms.\(^36–38\)

Children aged 8 to 10 years also completed the Terry to the YSR.\(^39,40\) a pictorial-based self-report that is a reliable and culturally sensitive version of the Dominic-R.\(^41,42\) The Dominic-R was designed to assess mental health symptoms and disorders in children 6 to 11 years of age by using established psychiatric criteria. It depicts a child named Dominic facing situations in the daily lives of children either in single pictures or in a set of drawings on the same page; sometimes Dominic is shown alone and sometimes with peers or adults. The interviewer reads aloud sentences about the visual stimulus that are also printed at the top and the bottom of the page so that the child who hears them can also read along. Because the Dominic-R was created and validated only for white children, the authors also developed and validated the Terry to depict a black character in identical situations. The situations and sentences are both positive (eg, Do you feel good at school, like Terry?) and negative (eg, Do you worry a lot about how you look, like Terry?) The child’s response “yes” or “no” is recorded. The manual provides criteria for determining presence of psychiatric disorder in 7 separate domains.

Children aged 11 to 12 years completed the Youth Self-Report (YSR) of externalizing and internalizing behaviors.\(^34\) The YSR, the child version of the CBCL, is a widely used instrument that provides data on a broad range of symptoms. It is appropriate for ages 11 through 18 years. It provides a total score, internalizing and externalizing subscales, and 8 syndrome scores. Psychometric data show reliabilities consistently above 0.80 in multiple data sets. The tool is sensitive and specific for identifying children and adolescents with clinically significant symptoms (CSS). We used the same criteria for the YSR as we used for the CBCL (ie, a total T score >63 [90th percentile] or a score above the 97th percentile on any individual subscale indicates CSS). Children who turned age 11 during the study were switched from the Terry to the YSR.

**Children’s Mental Health Service Use**

We assessed service use at each time period by asking mothers whether in the last 6 months the index child had received any counseling for a mental health or behavioral problem.

**Background Variables**

We assessed sociodemographic variables through maternal interview and included her age, race/ethnicity, birth place, educational level, employment, and source of income, as well as the child’s age and gender. We also asked mothers to report their use of marijuana/hashish, cocaine, crack, amphetamines/stimulants, inhalants (eg, glue, toluene, paint thinner), heroin, methadone (non-prescription), or any other drug in the past and in the last 6 months. Mothers were classified as never having used drugs, a past user, or a current user at baseline.

**Maternal Physical Health**

In the maternal interview, we obtained information on a range of physical health items, including time since the mother first learned about her HIV, her most recent T-cell count, whether she had experienced HIV-related health problems or opportunistic infections, whether she took protease inhibitors and/or drug “cocktails,” and numbers of hospitalizations and days spent in bed because of her health. We also asked whether she experienced limitations or restrictions on her work, housework, mobility or physical activity, or any other activity and whether she rated her health these days compared with others as excellent, very good, good, fair, or poor.

**Maternal Mental Health**

We used the Psychiatric Symptom Index (PSI)\(^42\) total score as a measure of maternal psychological distress at each of the study time points. The PSI used a 4-point scale (“never” to “very often”) to assess frequency of 29 common complaints or symptoms during the past 2 weeks. The PSI has excellent concurrent validity with criteria indicating psychological distress, such as help-seeking and use of psychoactive drugs.\(^42\) Although psychiatric diagnosis is not measured by the PSI, its items are consistent with established criteria used by clini-
cians. Moreover, in previous research with inner-city mothers, PSI total scores $>30$ had a sensitivity of 90% and a specificity of 58% when compared with a diagnosis of major depressive disorder obtained using a structured psychiatric interview. Therefore, we used this cutoff score to classify each mother as to whether she was likely to be clinically depressed at each assessment.

**Parent-Child Relationship**

The child’s interview included a revised version of the Inventory of Parent and Peer Attachment (IPPA), which is a self-report measure that assesses perceptions of the quality and security of attachments with parents and friends. We used only the 28 IPPA items that evaluated positive and negative aspects of the parent-child relationship. These items measure the degree to which the child trusts that the parent understands and respects his or her needs and is responsive and helpful with concerns. Examples of its items are, “When we talk about things, my mother listens to what I say,” and, “I get upset a lot more than my mother knows.” We reduced the number of response categories from 5 to 3 to make it easier for younger children to answer, because the original IPPA was designed to be administered to older adolescents (16–20 years). The response categories used were almost never, sometimes, and almost always.

**Data Analyses**

Each child mental health measure that we used has validated cutoff scores for CSS, which are types, frequency, or severity of emotional or behavioral symptoms that warrant evaluation and treatment by a clinician. We calculated whether the child had CSS on any measure at each time point. A “chronic” mental health problem was defined as CSS at $\geq 3$ of the 5 time points.

We used cross-tabulation and $\chi^2$ analyses to determine whether the percentages of children with CSS differed by child age or gender, mother’s drug use history, or assignment to Project Care or the control group. We also used $\chi^2$ analyses to examine the relationship of CSS to presence of chronic difficulties in maternal physical and mental health and in the parent-child relationship. For these analyses, mother’s physical health included the maternal self-rating of health status and her reported T-cell count. We categorized mother’s health rating as poor to fair versus good or better and her reported T-cell count as $<200$ vs $\geq 200$. Mother’s mental health was categorized as to whether the PSI total score was $\geq 30$, which we have noted is a likely indicator of depression. For parent-child relationship, we summed IPPA responses into a total score and then used a median split to dichotomize the sample at each time period, with scores above the median indicating a less close relationship. As with child’s clinically significant symptoms, a chronic problem in each area was defined as the mother’s having poor health, low T-cell count, likely depression, or poor parent-child relationship at $\geq 3$ of the 5 time points.

**RESULTS**

The demographic profile of parents in this study reflects the population of New York City families with HIV (Table 1): most mothers were black or Latina, 45% had not graduated from high school, and nearly three quarters had either a past or a present history of illicit drug use. The average child age was 10.5 years, and 58% were female.

Substantial proportions of mothers reported having problems or difficulties with their health in the year before the baseline survey. At enrollment into the study, the mean length of time since mothers had received their HIV diagnosis was 4.2 years (SD: 3.1), 35% reported that their most recent T-cell count was $<200$, and 31% said that they had been hospitalized at least once in the past year. In addition, 31% of mothers rated their health as fair or poor compared with others, 85% experienced $\geq 1$ activity restrictions because of their health, and 43% said that they had spent all or part of a day in bed in the last 2 weeks (mean number of days: 2.46). Moreover, 45% said that they had experienced HIV/AIDS-related health problems such as herpes, chronic yeast infections, thrush, candidiasis, or pneumonia. The majority were taking protease inhibitors (74%) and multiple HIV medications (77%).

Table 2 presents the percentage of children with CSS by measure over time. On every measure, more children had CSS at baseline, which is when maternal health data showed that the mothers were sickest. The highest number of children had CSS on the YSR and the CBCL at

<table>
<thead>
<tr>
<th>TABLE 1 Background Characteristics ($N = 157$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Maternal age, mean (SD), y</td>
</tr>
<tr>
<td>Maternal race/ethnicity, %</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Education, %</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Activity restrictions, %</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Drug use history, %</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Child age, mean (SD), y</td>
</tr>
<tr>
<td>Child gender, %</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
baseline. Overall, on any measure, 79% had CSS at baseline, 59% had CSS at 6 months, 52% had CSS at 12 months, 58% had CSS at 18 months, and 44% had CSS at 24 months.

During the 2-year study period, every (100%) child scored in the clinical range on at least 1 measure for at least 1 data point: 12% scored in the clinical range at 1 time point, 25% at 2 time points, 26% at 3 time points, 27% at 4 time points, and 9% at all 5 time points. Two thirds of children qualified for chronic CSS (3 of 5 time points). Chronic CSS were not significantly related to child age or gender, maternal drug use, or assignment to Project Care or the control group. Chronic CSS also were unrelated to mother’s health (ongoing poor health or low T-cell count), likely maternal depression, or a difficult parent-child relationship. Almost all children with CSS were recognized by both mother and child as having CSS. Only 8% of children were identified as having CSS by their own report and only 5% by their mother’s report only.

No 1 kind of emotional or behavioral problem was dominant. As Table 2 also shows, substantial proportions of children reported having high levels of depressive symptoms on the CDI during the study. We also found that among 8- to 10-years-olds, the most common psychiatric diagnosis in those who scored high on the Terry was separation anxiety. On the YSR, slightly more 11- to 12-year-olds scored above the clinical cutoff on internalizing symptoms than on externalizing symptoms at most assessments, and the highest proportions of CSS on this measure tended to be found on the somatization and social problems subscales. Compared with the children’s self-reports, parents rated more of these children as having high externalizing symptoms such as delinquency and aggression. However, among the internalizing sub-scales, parents also tended to rate more children above the clinical cutoff on somatization and social problems than on other CBCL syndrome subscales, which is similar to the pattern of child self-rating on the YSR.

Finally, we examined whether children had received counseling for a mental health or behavioral problem at any time during the 2-year study period. Two thirds of children received some mental health services during this period. Fewer than 25% received services at any 1 time period, and, over time, fewer children received care: 22% received care at baseline (79% had CSS), 19% at 6 months (59% had CSS), 23% at 12 months (52% had CSS), 16% at 18 months (58% had CSS), and 12% at 24 months (44% had CSS). Most children with chronic problems did receive mental health services, but 28% never saw a mental health professional.

**DISCUSSION**

Every school-aged child of a mother with late-stage HIV/AIDS had clinically significant psychiatric and/or behavioral symptoms during the 2-year study period. In this group of 8- to 12-year-old children whose mothers had HIV/AIDS, two thirds had chronic psychological problems (ie, persistent symptoms at least 3 of the 5 times they were measured). Most children were identified as having CSS by both themselves and their mother, although often they did not agree at the same time point or about the precise symptoms. Symptom scores were highest at baseline, when mothers were sickest, and declined over time. On the basis of previous analyses, we believe that the decline in symptoms among children is related to improved maternal physical health and mental health over time, although we cannot rule out the possibility that repeatedly measuring symptoms over time might have influenced responses.

The rate of psychological disorder found in this study is higher than in any other study that we found. Only one third of the general population of children studied cumulatively had significant mental health problems. Data on the mental health of other children at high risk, such as those who are in foster care, who are homeless, or who are poor, have not been reported longitudinally. Rates in available studies tended to be approximately half with a disorder, and some report rates as high as 8 in 10.

If a child’s mental health is assessed by parental or child report alone, then many children with serious psychological problems will be overlooked. Psychiatrists recommend that children be assessed by >1 reporter and that any reporter’s evaluation of symptoms be considered sufficient for assessment and referral. On the basis of this recommendation, children of mothers with HIV/AIDS should be evaluated by a parent as well as by their own self-report; teacher ratings are extremely valuable as well.

Moreover, children should be routinely screened for behavioral and psychological problems at different times.
using multiple measures. It is important to recognize that children may not exhibit problems at any 1 time or on any 1 tool. One reason is that children may have symptoms that are elevated but not sufficient to meet criteria for disturbance. More common, however, was that children who seemed to be in the reference range at 1 point in time had high symptom levels at another time. This can be attributable to a sudden new stress, such as a mother’s hospitalization, or to unpredictable or uncontrollable stressors, such as lack of social support, economic problems, or violence. It can also be attributable to the ongoing challenges of coping with chronic stress. Children in families in which a mother has HIV/AIDS are highly vulnerable, and any new change or threat is likely to tax their coping and resilience. Within just one 2-year window, a window in which parents were being treated successfully for HIV, every child experienced serious psychological disturbance. This sobering conclusion suggests vigilance and intensive assessment and referral.

The type of mental health problem that children experienced varied, although separation anxiety was most common in younger children, and internalizing problems were reported more by children than by their parents; parents reported more conduct problems, a finding consistent with the larger mental health literature. These children of HIV-infected mothers did not exhibit consistent profiles of symptoms across the sample; neither were symptoms consistent within individual children over time. This puts an additional burden on health professionals who are looking for ways to target clinical assistance to distressed children.

Many children did not receive consistent mental health care. Although most children received some care from a mental health worker, most of these were short-term therapies, and some were group-based, time-limited treatments. Many of those who received mental health services were seen by psychologists or social workers in their schools, who may have difficulty devoting intense, long-term assistance to any 1 child. Parents told us that when it was recognized that children needed professional assistance, most had to wait a long time to receive care, and when they finally did see someone, the high staff turnover resulted in children’s seeing multiple providers. Few mental health providers had specialized experience with HIV/AIDS and the challenges that these children were confronting at home. We cannot know from these data how many children would qualify for a diagnosis of psychiatric disorder and how many had elevated psychological symptoms that were appropriate to their situation. However, clinicians agree that a child who is in considerable emotional pain, even if contextually appropriate, should be cared for by a professional, both to help the child cope with the current stress and to initiate preventive interventions to reduce the likelihood that symptoms would worsen over time.

This study has several limitations. First, the sample consists of children in New York City, an urban epicenter of HIV/AIDS, at a time when antiretroviral therapy was just beginning. The findings should be generalized to different populations with great caution. Second, scores on the mental health measures are valid symptom reports, but they cannot be used to make diagnoses. Children in highly stressful situations, such as living with an ill mother, may demonstrate symptoms that are appropriate given their circumstances. The measures that we used are useful for identifying children whose psychological symptoms are elevated, but only a clinician can determine whether a given child warrants a clinical diagnosis. Third, we lack a control group and cannot interpret the burden of mental health symptoms in this population of children relative to similar children who live in an inner city and whose mothers do not have HIV/AIDS. However, as we noted previously, we did not identify a published study that reported rates of 100% disorder among a general population of children or among children in any high-risk group.

CONCLUSIONS

As the numbers of children who are living with a mother with HIV/AIDS increase, the numbers of children who need routine mental health assessment, referral, and treatment will increase as well. These children are hard to target for intervention because the marker for their risk is their mother’s health problem. Adult care providers may not take a family care perspective, and unless there is a system in place to identify and assist children of mothers with HIV/AIDS, they will continue to go unrecognized and untreated. Pediatricians are in a strategic position to assist children of mothers with HIV, because they often know the maternal health history, they see children over time in their practice, and they can identify and refer children with mental health problems to pediatric mental health services that can address their complex needs.

ACKNOWLEDGMENTS

This research was funded by National Institute of Mental Health grant R01MH55794 (to Dr Bauman).

We thank the field staff who recruited the participants and helped collect the data and the mothers and children who contributed their experiences.

REFERENCES

48. Okun A, Stein REK, Bauman LJ, Silver EJ. The content validity of the psychiatric symptom index. CES-depression scale and state-trait anxiety inventory from the perspective of DSM-IV. *Psychol Rep.* 1996;79:1059–1069
Out-of-Pocket Costs of Childhood Immunizations: A Comparison by Type of Insurance Plan

Noëlle-Angetérie M. Molinari, PhD, Maureen Kolasa, RN, MPH, Mark L. Messonnier, PhD, Richard A. Schieber, MD, MPH

Immunization Service Division, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

BACKGROUND. The “Guide to Community Preventive Services” strongly recommends reducing out-of-pocket costs to increase vaccination rates among children. Nevertheless, out-of-pocket expenses are still incurred during the receipt of childhood vaccines, vaccine administration, and associated well-child visits.

OBJECTIVE. Our goal was to estimate total and out-of-pocket costs of childhood immunization.

METHODS. We used the 2003 benefit-plan data for all 1217 private and public health plans registered in Georgia and the 2003 Advisory Committee on Immunization Practices recommended vaccine schedule to calculate costs to vaccinate children aged 0 to 5 years in 2003 dollars. By applying published estimates of health insurance enrollment of Georgia children, we calculated the total and out-of-pocket costs per child according to insurance status and race/ethnicity. Immunization coverage according to payer type was based on National Immunization Survey data.

RESULTS. Out-of-pocket costs ranged between $0 (Medicaid/Peachcare) and $652 (uninsured/Medicare). Most out-of-pocket costs were incurred during the first year of life. Up-to-date immunization status ranged from 63.7% for uninsured persons to 83.2% for privately insured persons. Up-to-date status was negatively correlated with out-of-pocket costs and the proportion of the population below 250% of the federal poverty level.

CONCLUSIONS. For most Georgia families, out-of-pocket expenses for childhood immunizations were low, favoring compliance with the recommended immunization schedule. However, families least able to afford the expense faced disproportionately high out-of-pocket costs. Out-of-pocket costs were inversely correlated with immunization coverage levels. Uninsured children whose families lived below 250% of the federal poverty level experienced the lowest immunization coverage levels. Immunization coverage through the Vaccines for Children Program and Medicaid/State Children’s Health Insurance Programs should be promoted to minimize or eliminate out-of-pocket costs related to childhood immunizations, especially among children of low-income families.
Reducing out-of-pocket (OOP) costs is an evidence-based strategy that is strongly recommended by the “Guide to Community Preventive Services” to increase vaccination rates among children.1–3 In 2003, the Advisory Committee on Immunization Practices recommended 23 immunizations for children before their fifth birthday, to be administered over 7 visits.4 Routine surveillance of 19- to 35-month-old children in the general US population in 2003 indicated that vaccination coverage for the 4:3:1:3:3 immunization series (≥4 doses of diphtheria, tetanus, and acellular pertussis vaccines, ≥3 doses of poliovirus vaccine, ≥1 dose of measles, mumps, and rubella vaccine, ≥3 doses of Haemophilus influenzae type b vaccine, and ≥3 doses of hepatitis B vaccine) was 79.4% (95% confidence interval [CI]: 78.5–80.3), with state means ranging between 67.5% and 94.0%.5

The total cost of each vaccine has 3 separate additive components: the cost of the vaccine itself, the cost of its administration by a health care provider, and the cost of a concomitant well-child examination. The US Vaccines for Children (VFC) program provides the vaccine free of charge to qualified children but does not cover the cost of vaccine administration or well-child examination.6 The latter 2 represent OOP expenses typically paid by the parent at the point of service. Although Medicaid-insured children face little or no OOP expense, privately insured children also incur OOP expenses in the form of copayments for 1 or more components.

The Task Force on Community Preventive Services first published its recommendation to reduce OOP costs for childhood immunizations in 1999.7 However, neither the studies systematically reviewed by the task force nor those published since then provide an estimate of OOP costs of the childhood vaccination schedule, for a general population or by payer type.8,9 Accordingly, our purpose was to (1) estimate OOP costs of childhood vaccination for children aged 0 to 5 years for a defined general population, Georgia’s 2003 birth cohort, stratified by type of insurance coverage and race/ethnicity and (2) determine the extent of any association between OOP and up-to-date (UTD) rates, controlling for race/ethnicity and poverty level. By detailing OOP costs for each payer type and quantifying the degree to which a reduction in OOP costs will increase immunization coverage for a defined general population, this study advances the topical literature.

METHODS

Insurance Status of Georgia’s 2003 Birth Cohort

The Current Population Survey Annual Social and Economic Supplement (CPS) 2004 data provides information on insurance status by age, race/ethnicity, income, and location. We used the CPS to calculate the number of children in Georgia’s 2003 birth cohort in each race and ethnicity group according to each payer category.9 Because the 2004 CPS reported insurance status for children younger than 5 years, we assumed that (1) the entire birth cohort survived to 5 years of age and (2) insurance distribution did not vary by age.

Calculation of Costs by Plan Type

This study incorporated data from multiple sources in an aggregate analysis that describes insurance and immunization coverage in Georgia (see Table 1 for details). We obtained benefit-plan data for 1217 health insurance plans registered in Georgia in 2003. These plans represented all potential payers in Georgia: (1) private insurers (employer-sponsored large-group plans, individually purchased and employer-sponsored small-group plans, public-sponsored vaccines, and Medicare-insured individuals), (2) Medicaid and Peachcare insured individuals, (3) military insured, and (4) children with private insurance.

<table>
<thead>
<tr>
<th>Database</th>
<th>No. of Health Plans</th>
<th>Population</th>
<th>Data Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medstat Marketscan (2003)</td>
<td>35</td>
<td>Large group, self-insured, employer sponsored, privately insured</td>
<td>Reimbursement and copay rates for well-child visit, vaccine administration, vaccine plan enrollment</td>
</tr>
<tr>
<td>Georgia Department of Insurance (2003)</td>
<td>1176</td>
<td>Individual and small group, privately insured</td>
<td>Copay rates for well-child visit, vaccine administration, vaccine plan enrollment</td>
</tr>
<tr>
<td>Georgia Department of Community Health (2003)</td>
<td>2</td>
<td>Medicaid and Peachcare insured</td>
<td>Reimbursement and copay rates for well-child visit, vaccine administration, vaccine plan enrollment</td>
</tr>
<tr>
<td>Tricare (2003)</td>
<td>3</td>
<td>Military insured</td>
<td>Reimbursement and copay rates for well-child visit, vaccine administration, vaccine plan enrollment</td>
</tr>
<tr>
<td>CDC Health Insurance Plan Enrollment Survey</td>
<td>1100</td>
<td>Individual and small group, privately insured</td>
<td>Plan enrollment for individually purchased and small-group employer-sponsored insurance</td>
</tr>
<tr>
<td>CPS (2004)</td>
<td>NA</td>
<td>Children aged 0–5 y in Georgia</td>
<td>Population estimates by insurance status/type and race/ethnicity</td>
</tr>
<tr>
<td>CDC vaccine price list</td>
<td>NA</td>
<td>Public- and private-sponsored vaccines</td>
<td>Vaccine prices for VFC and small-group private insurance</td>
</tr>
<tr>
<td>National Immunization Survey (2003)</td>
<td>NA</td>
<td>Children aged 19–35 mo in Georgia</td>
<td>Aggregate immunization coverage rates by insurance status and type</td>
</tr>
</tbody>
</table>

NA indicates not applicable.
and the military [Tricare]); (2) the uninsured; and (3) public payers (Medicare, Peachcare, Georgia’s State Children’s Health Insurance Program [SCHIP], and Medicaid). Table 2 describes the distribution of plans and the percentage of the 2003 Georgia birth cohort insured by each plan.9

For purposes of cost calculations, we assumed that the costs of services were fairly represented by the rate at which those services were reimbursed rather than some proportion of provider charges. We further assumed that (1) each child received all appropriate vaccines at the earliest eligible date according to the 2003 recommended immunization schedule,10 (2) each visit after the birth visit included a well-child examination, all age-appropriate vaccines, and a vaccine-administration fee for each injection, (3) all vaccines other than diphtheria, tetanus, and acellular pertussis and measles, mumps, and rubella were single-antigen vaccines, (4) no covered charges were disallowed by the insurer, (5) no vaccine shortages occurred, and (6) prices remained constant relative to each other during the study period. These assumptions reflect what parents could have reasonably expected when they anticipated the cost to vaccinate their child.

For each visit, the total cost per visit was calculated as the sum of the well-child examination fee, the vaccine antigen fee(s), and the vaccine-administration fee(s):

\[
\text{Cost/child} = \text{visit } 1 + \frac{\text{visit } 2}{1 + \frac{r}{m}} + \frac{\text{visit } 3}{1 + \frac{r}{m}} + \frac{\text{visit } 4}{1 + \frac{r}{m}} + \cdots + \frac{\text{visit } 7}{1 + \frac{r}{m}}^{10}
\]

(1)

OOP costs per visit were calculated by applying any coinsurance or copayments to the fees charged for the well-child visit, the vaccine(s), and the vaccine-administration fee(s). For example, suppose a privately insured 6-month-old child received a well-child examination (fee = $75) and 5 vaccines totaling $175 in cost, with vaccine-administration fees of $10 for the first vaccine and $15 for each additional vaccine. The child would incur a total cost for the visit of $75 + $175 + $10 + $60, or $320. If insurance covered all costs other than a $20 copayment for the visit and 10% coinsurance for the vaccines and their administration, the total OOP costs would equal $20 + ($20 × 0.10) = $24.50, or a total of $45.50 for that visit.

Reimbursement rates, copayments, and coinsurance rates for vaccines, vaccine-administration fees, and well-child examinations were obtained from separate sources, which depended on the type of insurer. For large-group, employer-sponsored private insurance plans, we obtained average reimbursement rates, copayments, and coinsurance rates for vaccine antigens, vaccine-administration fees, and well-child fees from Medstat’s Marketscan 2003 insurance claims database according to plan type in Georgia.11 For individual and small-group private insurance plans, we obtained average reimbursement rates according to plan type in Georgia for vaccine-administration fees and well-child fees from Medstat’s Marketscan 2003 insurance claims database. Private-market wholesale vaccine prices were obtained from the Centers for Disease Control and Prevention (CDC) vaccine price list for 2003 and were increased by 25% to adjust for retail pricing.12,13 Using benefit-plan data filed with Georgia’s Department of Insurance in 2003, we obtained copayments and coinsurance rates for individual and small-group private insurance plans. We calculated average OOP and total costs per visit for each plan type, and results were weighted to reflect plan type distribution by using data from the CDC’s Health Insurance Plan Enrollment Survey (Tricare South, unpublished data, 2005).

For children served by military health insurance (Tricare), the average total cost per visit was calculated by using military reimbursement rules on allowable reimbursements for vaccine antigens, vaccine administration, and well-child examinations in Atlanta and the rest of Georgia. Average OOP costs per visit were based on Tricare Extra copayments and coinsurance rates (Georgia Department of Community Health, unpublished data, 2004). Results were weighted by using the proportion of Georgia medical treatment facilities located in the Atlanta area versus the rest of Georgia to reflect military enrollments in the state.

In Georgia, children who are covered by Medicaid, Peachcare, and Medicare and those who are uninsured qualify for free vaccines through the VFC program.14 For children who are covered by Medicaid and Peachcare, Georgia Medicaid regulations governed reimbursement rates, vaccine-administration fees, and well-child visit rates for both Medicaid and Peachcare providers. Medicaid charged no monthly premium. Infants from households with incomes up to 200% of the federal poverty level (FPL) and children aged 1 to 5 years from households up to 133% of the FPL were eligible for Medicaid. The Peachcare monthly premium was graduated on the basis of the ability to pay and did not exceed $20 for ≥2 children in the same household in 2003. Children from households up to 235% of the FPL were eligible for the program.14 We used data on Medicaid-allowable charges and VFC vaccine prices to calculate the total costs per visit for Medicaid- and Peachcare-covered children.12,14 Medicare did not cover well-child examinations or child-

*We concluded that well-child visits with each immunization after birth because this represents the gold standard of care. Although uninsured children can receive immunizations at a federal qualified health center or rural health center without the concurrent well-child visit, this represents a suboptimal outcome. Furthermore, pediatricians do not routinely provide immunizations without the concurrent well-child visit.
hood vaccines in 2003. Therefore, Medicare-covered children were eligible for free vaccines through the VFC but had no insurance coverage for vaccine-administration fees and well-child visits. As a result, Medicare-covered children faced the same OOP costs for immunizations as did children with no health insurance coverage. Although the VFC does not cover vaccine-administration charges, the maximum vaccine-administration fee that VFC providers are allowed to charge is set by the Centers for Medicare and Medicaid Services. We used VFC vaccine prices, Georgia VFC-allowed vaccine-administration charges, and Medicaid-allowable charges for well-child examinations to calculate cost per visit for Medicare-covered and uninsured children.

To calculate the cost in 2003 dollars of a child’s vaccinations and associated well-child visits, we summed the cost of the 7 visits and then discounted on a monthly basis by using a 3% annual discount rate:

\[
\text{Total Cost} = \text{Well-Child Exam Fee} + \text{Vaccine Antigen Fee} + N_{\text{Vax}}(\text{VaxAd min Fee})
\]

Here, visit 1 represents the cost of the first visit, \( r \) is the annual discount rate of 3%, and \( m \) is 12, which represents the monthly basis for compounding. We performed a similar calculation to obtain the present value of OOP costs of a child’s vaccinations and associated well-child visits for children aged 0 to 4 years.

We multiplied the total and OOP costs per child in each payer category by the number of children in each payer category. Our results represent the population-weighted total and OOP costs to immunize Georgia’s 2003 birth cohort.

### Relationship Between OOP Costs and Immunization Coverage

To assess possible associations between costs and UTD coverage, we calculated the correlation between (1) 4:3:1:3:3 UTD immunization coverage and the 4:3:1:3:3 UTD immunization coverage and the proportion of population below 250% of the FPL by using the National Immunization Survey 2003 data. Statistical analysis was conducted by using Pearson correlation coefficients and simulation by using a weighted generalized linear model with log link, which was weighted for the proportion of population in each insurance category and given race/ethnicity in that insurance category. Keeping in mind that the 7 payer types represented the state of Georgia, we bootstrapped our results by (1) randomly sampling 1000 times from the data with replacement, assuming a log-normal distribution and (2) generating data from a fitted distribution for simulation.

### RESULTS

In the Georgia 2003 birth cohort, most children (66%) were covered by private insurance (57% large-group insurance, 7% small-group/individual insurance, and 2% military). Almost 14% were not covered by any type of insurance. The remaining 20% of children were publicly insured (1% Medicare, 8% Peachcare, and 12% Medicaid). Among privately insured children, ~61% were white, 34% nonwhite, and the remaining 5% Hispanic. Nonwhite children were primarily black but also included children of American Indian, Alaskan Native, Native Hawaiian, Pacific Islander, Asian, and mixed-race descent. Uninsured children were 34% white, 45% nonwhite, and 21% Hispanic. Children covered by public insurance were 51% white, 40% nonwhite, and 8% Hispanic. Table 2 shows the Georgia 2003 birth cohort according to insurance plan type and race/ethnicity.

During the first year of life, the typical immunization visit, which occurs after the birth visit, cost an average of $264 (range: $192–$347). OOP costs averaged $59 per
visit (range: $0–$122). When considering all immunization visits of children through 4 years of age, after the birth visit, the average visit cost was $228 (range: $166–$297), and OOP costs averaged $53 per visit (range: $0–$111). Children covered by Medicaid/Peachcare had the lowest average cost per immunization visit and lowest average OOP cost. Children covered by individual/small-group insurance had the highest average cost per visit. Uninsured children and those covered by Medicare had the highest average OOP cost per visit. Table 3 presents total and OOP costs for each immunization visit according to insurance status and type and total costs to vaccinate a child.

In general, vaccines accounted for the largest portion of total cost, followed by well-child visits and administration fees (Fig 1). In contrast, vaccines represented the smallest component of OOP costs, followed by administration fees and well-child visits. Total costs were highest for the large-group privately insured and declined toward public insurance. OOP costs were lowest for Medicaid/Peachcare, which was followed by the large-group privately insured. OOP costs were highest among the uninsured and children covered under Medicare. Approximately 13% of total costs and 0% of OOP costs were attributed to this group.

More than 55% of Georgia’s 2003 birth cohort was white; ~57% of total costs and 48% of OOP costs were attributed to them. Nearly 37% of the cohort was nonwhite; almost 36% of total costs and 39% of OOP costs were attributed to them. Although just over 7% of the birth cohort was white Hispanic, ~7% of total costs and 13% of OOP costs were attributed to these children. The remaining 0.6% of the cohort was nonwhite Hispanic; they paid ~0.5% of total costs and 0.4% of OOP costs.

By using simple 1-way correlation, immunization coverage was negatively correlated with OOP costs to bring children UTD for 4:3:1:3:3 ($p = -0.71$) and negatively correlated with the proportion of enrollees below 250% of the FPL ($p = -0.49$) (Tables 4 and 5). When controlling for race/ethnicity and the proportion of the cohort in each insurance category, 1% increase in OOP cost of the 4:3:1:3:3 series was associated with a 0.07% (95% CI: 0.090 to 0.056; $p < .01$) reduction in

### TABLE 3 Costs of Immunization and Well-Child Visits Through 4 Years of Age

<table>
<thead>
<tr>
<th>Payer Category</th>
<th>Visit 1 (Birth)</th>
<th>Visit 2 (2 mo)</th>
<th>Visit 3 (4 mo)</th>
<th>Visit 4 (6 mo)</th>
<th>Visit 5 (12 mo)</th>
<th>Average (First Year)</th>
<th>Total (First Year)</th>
<th>Visit 6 (15 mo)</th>
<th>Visit 7 (48 mo)</th>
<th>Average (4 y)</th>
<th>PV Total (4 y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>369</td>
<td>303</td>
<td>355</td>
<td>359</td>
<td>347</td>
<td>1431</td>
<td>117</td>
<td>231</td>
<td>297</td>
<td>1729</td>
</tr>
<tr>
<td>OOP</td>
<td>1</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>80</td>
<td>19</td>
<td>19</td>
<td>20</td>
<td>115</td>
</tr>
<tr>
<td>Individual/small</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>362</td>
<td>299</td>
<td>345</td>
<td>373</td>
<td>345</td>
<td>1420</td>
<td>123</td>
<td>230</td>
<td>297</td>
<td>1722</td>
</tr>
<tr>
<td>OOP</td>
<td>5</td>
<td>36</td>
<td>33</td>
<td>36</td>
<td>37</td>
<td>35</td>
<td>146</td>
<td>22</td>
<td>28</td>
<td>32</td>
<td>190</td>
</tr>
<tr>
<td>Military</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
<td>279</td>
<td>224</td>
<td>256</td>
<td>283</td>
<td>260</td>
<td>1090</td>
<td>111</td>
<td>179</td>
<td>227</td>
<td>1345</td>
</tr>
<tr>
<td>OOP</td>
<td>7</td>
<td>56</td>
<td>45</td>
<td>51</td>
<td>57</td>
<td>52</td>
<td>218</td>
<td>22</td>
<td>36</td>
<td>45</td>
<td>269</td>
</tr>
<tr>
<td>Uninsured</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>222</td>
<td>198</td>
<td>222</td>
<td>239</td>
<td>220</td>
<td>906</td>
<td>83</td>
<td>139</td>
<td>189</td>
<td>1096</td>
</tr>
<tr>
<td>OOP</td>
<td>15</td>
<td>129</td>
<td>115</td>
<td>129</td>
<td>115</td>
<td>122</td>
<td>503</td>
<td>70</td>
<td>100</td>
<td>111</td>
<td>652</td>
</tr>
<tr>
<td>Public</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>222</td>
<td>198</td>
<td>222</td>
<td>239</td>
<td>220</td>
<td>906</td>
<td>83</td>
<td>139</td>
<td>189</td>
<td>1096</td>
</tr>
<tr>
<td>OOP</td>
<td>15</td>
<td>129</td>
<td>115</td>
<td>129</td>
<td>115</td>
<td>122</td>
<td>503</td>
<td>70</td>
<td>100</td>
<td>111</td>
<td>652</td>
</tr>
<tr>
<td>Medicaid/Peachcare</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>190</td>
<td>173</td>
<td>190</td>
<td>214</td>
<td>192</td>
<td>784</td>
<td>78</td>
<td>123</td>
<td>166</td>
<td>957</td>
</tr>
<tr>
<td>OOP</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

All amounts are shown in 2003 US dollars. PV indicates present value.
4:3:1:3:3 UTD coverage. Similarly, a 1% increase in the proportion of enrollees below 250% of the FPL was associated with a 0.09% (95% CI: −0.105 to −0.071; \( P < .05 \)) reduction in 4:3:1:3:3 UTD coverage.

**DISCUSSION**

This study, an aggregate analysis that incorporated data from multiple sources that describe insurance and immunization coverage in Georgia, indicated that ~12% of all vaccination costs were paid OOP. The burden of these OOP costs fell disproportionately on the uninsured; although the uninsured comprised 14% of the population, they paid >50% of all OOP costs. These children also experienced the lowest UTD immunization rates of children in our sample (63.7%). Note that 70% of Georgia’s uninsured children lived in households with an income below 250% of the FPL; at least 1 adult was employed in 78% of the uninsured households. In contrast to the uninsured, children covered by Medicaid and Peachcare experienced the lowest OOP costs, and immunization coverage among this group was 79.1%, which was nearly equivalent to coverage among children who were privately insured. This finding, that low OOP costs are associated with increased coverage, indicates the value of decreasing OOP costs for low-income families in promoting immunization coverage.

A number of studies have suggested that decreasing OOP costs will increase immunization coverage. The Task Force on Community Preventive Services noted in their “Guide to Community Preventive Services” that strong evidence indicates the effectiveness of increasing UTD coverage by reducing OOP costs. In their review, a 15% median increase in vaccination coverage was achieved by a reduction in OOP expenses, by providing free vaccines, reducing administrative costs, providing insurance when it was lacking, reducing copayments at.

---

**FIGURE 1**

Total and OOP costs (2003 US dollars) per child according to insurance type for children aged 0 to 5 years.

**FIGURE 2**

Total and OOP costs (2003 US dollars) per child according to race/ethnicity for children aged 0 to 5 years. Nonwhite includes black, Asian, American Indian/Alaskan Native, Native Hawaiian/Pacific Islander, and ≥2 races.
the point of service. Our study quantified the degree to which a reduction in OOP cost could be effective, and we found that a 1% decrease in OOP cost would result in a 0.07% increase in UTD coverage. Assuming constant elasticity, the elimination of all OOP expenses in Georgia (ie, a reduction of the mean OOP costs of $175 per person to $0) would result in a 7% UTD increase in immunization coverage for the 4:3:1:3:3 series, which would bring coverage to 85.4% statewide at a cost of $19.1 million.

One strategy to decrease OOP expenses for childhood immunizations is to increase enrollments in Medicaid/SCHIP of the uninsured and underinsured who are already eligible and/or to relax eligibility requirements. However, increased enrollment would have budget implications beyond immunization-related costs and the funding of both Medicaid and SCHIP have not kept pace with enrollment expansion.25–30 In fact, in 2005, 20 states tightened Medicaid/SCHIP eligibility criteria, which made enrollment in the programs more difficult, and 21 states reduced program benefits and/or increased cost sharing.26 These measures may contribute to reduced immunization coverage among those not eligible for enrollment. A sustained, comprehensive, state-based approach to these regulations is needed to reverse this trend.

In our study, the majority of OOP immunization costs were incurred during the first year of life, when parents often experience a reduced income that is related to time off work for the mother and additional expenses for the new child. Parents who anticipate preventive care consider OOP vaccination costs against expenses for other necessities. Annually, the average family of 4 spends $2581 for health care and $7472 for food.31 During the first year of a child’s life, average OOP vaccination costs are $242. Although these OOP vaccination costs may not seem steep by comparison, for the uninsured, average OOP vaccination costs during the first year are $652. These OOP costs may compete unsuccessfully with other monthly necessities, especially because, in 2003, 11% of families in Georgia lived below the FPL ($18 400 annually for a family of 4).31

Although children covered by Medicare represented only a very small proportion of children in Georgia (0.4%), OOP costs experienced among these children on an individual level were equivalent to those incurred by uninsured children. Children covered by Medicare are dependents of Medicare beneficiaries, such as children in the care of grandparents, who may experience unique difficulty in covering immunization-related costs because of their fixed income and their own increasing age-related personal medical costs. These children should be considered to be at risk for underimmunization. Medicare coverage of well-child visits and administrative fees associated with dependent minor children’s immunization visits would help ensure that such children receive recommended immunizations. To help ensure that such children receive recommended immunizations, the well-child visit and administrative fees associated with dependent minor children’s immunization visits should be covered by Medicare.

OOP costs per child did not differ significantly according to racial/ethnic group when controlling for payer category and poverty, except for white Hispanic children. White Hispanic children, who make up 21% of the Georgia’s uninsured children but represent only 14% of the total population, incurred the highest OOP costs.

This study has several limitations. The analysis was conducted at the aggregated insurance-plan type level and may, therefore, be subject to ecologic fallacy. However, the negative association between OOP price and demand for a product has been well documented at the individual and aggregate levels, which suggests that ecologic fallacy is unlikely. The analysis was limited to Georgia, which is not representative of the United States. Georgia, however, does include a demographically diverse population, and Peachcare, Georgia’s SCHIP, is separate from Medicaid, as is the case in 18 other states. The Health Insurance Plan Enrollment Survey (CDC institutional review board protocol 4205), which provided plan enrollments used to weight private insurance, had a 61% response rate. With the exception of large-group insurance, copayments were based on benefit-plan descriptions. Weights for military insurance were based on the locations of medical treatment facilities rather than the locations of military personnel in Georgia. Although we do not believe these limitations greatly impacted our results, the degree of sensitivity of the

### TABLE 4 OOP Payment Rate and UTD Percentage of 4:3:1:3:3

<table>
<thead>
<tr>
<th>Payer Category</th>
<th>OOP Rate, %</th>
<th>4:3:1:3:3 UTD Percentagea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employer sponsored/individual and small group</td>
<td>7</td>
<td>83.2</td>
</tr>
<tr>
<td>Military</td>
<td>20</td>
<td>80.2</td>
</tr>
<tr>
<td>Uninsured</td>
<td>60</td>
<td>63.7</td>
</tr>
<tr>
<td>Public</td>
<td>0</td>
<td>79.1</td>
</tr>
</tbody>
</table>

a UTD Percentage indicates the percent of 19- to 35-month-old children who are UTD for the recommended immunizations in the 4:3:1:3 series.

### TABLE 5 Correlates for UTD Percentage of 4:3:1:3:3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Correlation Coefficient</th>
<th>Elasticityb</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>OOP cost of 4:3:1:3:3a</td>
<td>−0.71</td>
<td>−0.09 to −0.06</td>
<td></td>
</tr>
<tr>
<td>Proportion at &lt;250% of the FPLb</td>
<td>−0.49</td>
<td>−0.10 to −0.07</td>
<td></td>
</tr>
</tbody>
</table>

Note that correlation coefficient and elasticity are unit-free measures.

a Significant at <1%

b Significant at <5%

c Controls for race/ethnicity and percent population in each insurance category.
results to these assumptions is reflected in the CLs surrounding our estimates.

CONCLUSIONS
For most Georgia families, OOP expenses for childhood immunizations were fairly low, which favored compliance with the recommended immunization schedule. However, families who were less able to afford the expense, the uninsured, faced disproportionately high OOP costs. OOP costs were significantly inversely correlated with immunization coverage levels so that uninsured children whose families lived below 250% of the FPL experienced the lowest immunization coverage levels. Enrollment in and immunization insurance coverage through the VFC program and Medicaid/SCHIP should be promoted to minimize or eliminate OOP costs related to childhood immunizations, especially among children of low-income families.

ACKNOWLEDGMENTS
We thank Dr Philip J. Smith, who provided Georgia's immunization coverage rates, and Edith Gary and Heather Purk for research assistance.

REFERENCES
8. Joyce T, Racine A. CHIP shots: association between the State Children's Health Insurance Programs and immunization coverage and delivery. Pediatrics. 2005;115(5). Available at: www.pediatrics.org/cgi/content/full/115/5/e526
Obesity Is a Common Comorbidity in Children With Congenital and Acquired Heart Disease

Nelangi M. Pinto, MD, Bradley S. Marino, MD, MPP, MSCE, Gil Wernovsky, MD, Sarah D. de Ferranti, MD, MPH, Amy Z. Walsh, RN, Meena Laronde, RN, Kristen Hyland, BA, Stanley O. Dunn, Jr, BA, Meryl S. Cohen, MD

Division of Cardiology, Department of Pediatrics, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania; Department of Cardiology, Children’s Hospital Boston, Boston, Massachusetts

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVES. Obesity may pose additional cardiovascular risk to children with acquired and congenital heart disease. Many children with heart disease are sedentary as a result of physician-, parent-, and/or self-imposed restrictions. The aim of this study was to evaluate the impact of the epidemic of obesity on children with heart disease.

PATIENTS AND METHODS. A cross-sectional review was performed of children evaluated in 2004 at 2 cardiology outpatient clinics. Differences in the prevalence of obese (BMI ≥ 95%) and overweight (BMI 85%–95%) children were compared with national data and healthy control subjects. Dictated letters were reviewed to determine whether obesity was discussed with referring practitioners.

RESULTS. Of 2921 patients assessed, 1523 had heart disease. Diagnostic subgroups included “mild” heart disease (n = 401), arrhythmia (n = 447), biventricular repair (n = 511), univentricular palliation (Fontan; n = 108), and heart transplantation (n = 56). More than 25% of the patients with heart disease were obese or overweight; the prevalence of obese and overweight children was significantly lower only in the Fontan group (15.9%). Pediatric cardiologists failed to document obesity or weight counseling in the majority of clinic letters.

CONCLUSIONS. Obesity is common in children with congenital and acquired heart disease. Pediatric cardiologists demonstrate inadequate communication regarding this problem to referring practitioners. Healthy-lifestyle counseling and routine exercise in children with heart disease may be underemphasized.
CHILDHOOD OBESITY HAS increased dramatically in the last half century. The most recent data published by the Center for Disease Control’s serial National Health and Nutrition Examination Survey (NHANES) indicates a 45% increase in the prevalence of obese children in the past decade.1 Childhood obesity has been shown to be a significant predictor of obesity in adulthood.2,3 Furthermore, childhood obesity is associated with systemic hypertension, type 2 diabetes, and endothelial dysfunction and is an independent predictor of coronary artery disease and premature death in adults.4–7 Obese adolescents exhibit changes in left ventricular mass related to an increased cardiac workload.8 Even with normal ventricular mass, overweight children exhibit subtle changes in diastolic function that may have implications for their future cardiovascular health.9

The impact of obesity on long-term cardiovascular health may be particularly important to children with preexisting congenital or acquired heart disease. Although much emphasis is placed on preventive medicine in adult cardiology, it is not often a focus of the pediatric cardiologist outpatient visit despite recent guidelines on protecting future cardiovascular health in childhood.10 Children with congenital heart disease may have unique risk factors that contribute to the development of obesity. Children with certain types of heart disease may be at risk for failure to thrive in early infancy. Thus, practitioners have historically focused on encouraging weight gain, particularly in early childhood, when long-term eating habits begin to develop. Moreover, children with some forms of heart disease are commonly cautioned against certain types of physical activity. Recently, it has been shown that formal activity restriction in children with heart disease may be a predictor of the development of obesity.11

Very limited data exist about the prevalence of obesity in children with congenital and acquired heart disease. Because survival is now expected into adulthood for many of these patients,12 identifying the degree to which this additional cardiovascular risk factor exists within this population is likely to have important implications for long-term outcomes. The purpose of this study was to estimate the prevalence of obesity and overweight in children with heart disease, to assess the impact of obesity on systolic blood pressure measurement in this population, and to determine whether weight counseling was documented in referral correspondence to primary care physicians.

PATIENTS AND METHODS
Charts from outpatient cardiology clinic visits at 2 centers, Children’s Hospital of Philadelphia (CHOP) and Children’s Hospital Boston (CHB), for the calendar year 2004 were reviewed. All of the patients aged 6 to 19 years at the time of their visit were candidates for inclusion (the same age group highlighted by the National Center for Health Statistics from the 1999–2002 NHANES).2 Exclusion criteria included factors independently associated with obesity or failure to thrive, such as a known genetic syndrome, mechanical support (such as a tracheostomy or gastrostomy tube), a diagnosis of protein-losing enteropathy, or referral specifically for hyperlipidemia or obesity. Systematic selection of patients was achieved by sampling every third patient by date of visit at CHOP and by sampling every third patient by alphabetized last name at CHB. Patients with multiple visits during the year had only their most recent clinic visit included.

Cardiology outpatient letters are the equivalent of clinic notes at CHOP and CHB. Data collected from cardiology outpatient letters included gender, age at visit, weight, height, heart rate, blood pressure, primary cardiac diagnosis, type of surgical or catheter intervention if performed, and documentation of discussion of weight. Documented blood pressure data were 1-time measurements obtained with either a Dinamap automated blood pressure monitor or by manual sphygmomanometer at the time of the clinic visit. Patients were grouped into 5 diagnostic heart disease categories: (1) “mild heart disease” that has not been treated with either surgical or catheter intervention (eg, mild valvar disease, hemodynamically insignificant septal defects, or Kawasaki’s disease); (2) primary rhythm disturbances in patients with structurally normal hearts (electrophysiology [EP]); (3) congenital heart disease with surgical and/or catheter-directed intervention culminating in a biventricular circulation (biventricular repair); (4) complex congenital heart disease (typically a univentricular heart) resulting in Fontan palliation (Fontan); and (5) cardiac transplantation. “Clinic control subjects” were healthy patients referred to the 2 outpatient clinics and discharged with a diagnosis of no heart disease (eg, chest pain, palpitations, or functional murmurs).

BMI was calculated from weight and height data and then plotted on Centers for Disease Control and Prevention BMI curves to determine age- and gender-appropriate percentiles. For the purposes of this study, “obese” was defined as BMI =95th percentile and “ overweight” was defined as BMI in the 85th to 95th percentile.13 Blood pressure measurements were similarly converted to percentiles that were standardized for age, gender, and height.14

Statistical Analysis
Descriptive data were expressed as means with SDs (normally distributed continuous variables), as medians and ranges (nonnormally distributed continuous data), or as proportions (categorical and dichotomous variables). Continuous variables were evaluated for central tendency and variability with the Shapiro-Francia test. Systolic blood pressure percentiles were graphed by box-and-whisker plots. Diastolic blood pressure data were
not analyzed, because they are inherently unreliable by automated measurement.

The Wilcoxon rank-sum test was used to compare differences in age and BMI between control subjects and patients with heart disease. Fisher’s exact test was used to compare dichotomous variables, including gender, the presence of obesity and heart disease, diagnostic categories, and documentation of weight counseling. The Wilcoxon rank-sum test was also used to determine whether there was any difference in age or blood pressure percentile based on overweight category.

Obesity and overweight proportions for the total cohort, each center, the heart disease group, and individual diagnostic categories were obtained from contingency tables. The prevalence of obesity and overweight were compared against the Centers for Disease Control and Prevention NHANES control data, cohort patients without heart disease who were treated as internal control subjects, and between the 2 centers using the 2-sample test of binomial proportions. This test was also used to compare weight counseling between centers.

RESULTS
After selection, 2921 patients met criteria for this study. Study population characteristics are described in Table 1. There was no difference in age or gender between the clinic control subjects and patients with heart disease. Table 2 describes the breakdown of patients at each center. Patients without heart disease (clinic control subjects; \( n = 1398 \) [47.9%]) made up the largest percentage of the study population, both in total and at each center. A significantly higher percentage of patients with mild heart disease and biventricular repair were selected from CHB, whereas more EP patients were selected from CHOP. Table 3 describes patients with heart disease from both centers. There was no difference in age between the patients with heart disease who were obese and overweight and those with a normal BMI.

The overall prevalence of obesity in the total cohort was 15.2%, and combined obesity and overweight was 29.1%. This was not different from the Centers for Disease Control and Prevention NHANES national prevalence data for children aged 6 to 19 years of 16% obese and 31% either obese or overweight (Fig 1A). Healthy patients without heart disease (clinic control subjects) were no different in their prevalence of obesity and overweight than those in the NHANES population, suggesting that the 2 outpatient control populations were representative of these national data. Patients with any type of heart disease (combined assessment of mild heart disease, biventricular repair, Fontan, EP, and transplant groups) had a lower prevalence of obese (13.8%) and combined obese and overweight patients (26.2%) than the healthy clinic controls (\( P = .032 \) and \( P < .001 \), respectively). When patients with heart disease were divided into diagnostic categories, transplant, EP, and mild heart disease groups had a similar prevalence of obesity and overweight as the healthy control population (Fig 1B). Patients with palliated congenital heart disease, both biventricular repair and Fontan, had a lower prevalence of obesity and overweight (Fig 1B). Although the prevalence was statistically lower than in the healthy control patients, 23.9% of patients with a biventricular repair and 15.9% of Fontan patients were obese or overweight. Healthy control patients from CHB had a statistically lower prevalence of obesity than those seen at CHOP, but the prevalence of obesity in children with heart disease was similar at both centers. In addition, there was no difference in the prevalence of overweight between the centers in either the healthy control or heart disease patients.

Figure 2 compares the systolic blood pressure percentiles for normal weight, overweight, and obese patients (unadjusted for medication use). For those with any form of heart disease, the systolic blood pressure percentile was significantly higher in the obese and overweight groups compared with those with normal BMI (\( P < .001 \)). This trend was also shown in the healthy clinic control subjects.

Discussion of weight, as documented by clinic letters,
occurred in only 13.3% of obese patients and 8.7% of overweight patients who were followed for acquired and congenital heart disease. There was no statistically significant difference between the frequency of documented weight discussion between CHOP and CHB.

**DISCUSSION**

Similar to the alarming trend of obesity in otherwise healthy children, our study shows that there is a disturbingly high prevalence of obesity and overweight in children with the additional risk factors of congenital or acquired heart disease. More than one quarter of children with heart disease are obese or overweight.

The impact of obesity on other pediatric patient populations with chronic diseases, such as cystic fibrosis, end-stage renal disease, and acute myeloid leukemia, has been reported previously. To our knowledge, there are no previous large studies that look at the scope of obesity in pediatric patients with heart disease from multiple geographic regions. The current study, including >1500 patients with acquired or congenital heart disease, allowed for comparison among different categories of pediatric heart disease. The prevalence rates of obesity were similar across all of these groups, with the exception of patients with the most complex heart disease who have undergone the Fontan procedure. This is

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Normal BMI</th>
<th>Overweight</th>
<th>Obese</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (range), y</td>
<td>12.0 (6.0–19.0)</td>
<td>13.0 (6.0–19.0)</td>
<td>11.0 (6.0–19.0)</td>
<td>.90/26</td>
</tr>
<tr>
<td>Systolic BP, median (range), %</td>
<td>62.10 (0.04–100.00)</td>
<td>75.70 (0.06–99.90)</td>
<td>79.90 (0.95–100.00)</td>
<td>&lt;.001b</td>
</tr>
</tbody>
</table>

BP indicates blood pressure

* P values are 2-sided, normal versus overweight/normal versus obese.

**TABLE 3**

Selected Variables According to BMI Category for Patients With Heart Disease

*FIGURE 1*

Prevalence of obesity and overweight. A, Prevalence of national data (from NHANES 1999–2002) compared with the clinic cohort and patients with heart disease (HD). The heart-disease group includes all patients who had mild heart disease, EP, biventricular repair (2V), univentricular palliation, and transplant (see text for details). * P < .03 when compared with control patients. B, Prevalence of national data and clinic controls against categories of heart disease. ** P < .01 for comparisons with national data and controls. The P values not given were all not significant.
not surprising, because patients with univentricular hearts often have chronic, long-term issues, including heart failure and failure to thrive.\textsuperscript{18}

Although the health risks associated with obesity in the general pediatric population are well established,\textsuperscript{19,20} the truly long-term outcomes for pediatric patients with heart disease await longer follow-up. Life span for children with congenital heart disease is less than normal.\textsuperscript{21,22} Congenital and acquired coronary disease and underlying vascular abnormalities in childhood are likely to be exacerbated by obesity, which is independently associated with endothelial dysfunction and hypertension.\textsuperscript{20,23} In particular, certain populations, such as those with congenital abnormalities of coronary blood flow,\textsuperscript{24,25} acquired abnormalities, such as graft vasculopathy after cardiac transplantation\textsuperscript{26} or Kawasaki disease,\textsuperscript{27} or those who have had surgical manipulation of the coronary arteries (eg, the arterial switch or Ross operations),\textsuperscript{28,29} have been shown to have abnormalities in coronary function in later childhood and may be at increased risk for atherosclerosis as adults. This risk may also be exacerbated in children with previous repair of coarctation of the aorta where impairment of endothelial function and increased carotid intimal thickness have
been identified. Although the long-term impact of superimposed obesity on congenital and acquired heart disease remains unknown, it is likely to increase morbidity and mortality as it does in adults with heart disease.

Previous studies have detailed relationships among blood pressure, hypertension, and obesity. A recent study of a primary pediatric outpatient population found that obese and overweight children had higher systolic and diastolic blood pressures than normal-weight patients. Our study found that obese and overweight patients with heart disease were significantly more likely to have increased systolic blood pressure percentiles during 1-time measurement in the outpatient setting. Although this study was not designed to evaluate hypertension in this population, the trend toward higher blood pressure in obese patients raises concern. In children with congenital heart disease, hypertension may not only impact their long-term cardiovascular health but may also have more immediate implications for their underlying disease (such as valvar and ventricular dysfunction). Given the severity of univentricular heart disease, obesity and overweight (in 15.9% of the Fontan group) in this high-risk population are likely to have a negative impact secondary to the risks associated with increased afterload and ventricular mass.

Despite a significant proportion of children who are obese and overweight in both outpatient practices, cardiovascular practitioners at both centers failed to document obesity and/or weight counseling in >85% of these patients with heart disease. Although it is unknown whether such a discussion took place during the visit itself, lack of documentation to the referring physician in the clinic letter and/or note suggests that communication regarding the problem of obesity in association with heart disease did not occur with the primary practitioner. As specialists, pediatric cardiologists tend to focus their clinic visit on the primary reason for referral. In addition, they may believe that their patients and parents will not be receptive to weight counseling at their “heart” clinic visit. Pediatric cardiologists have historically focused on adequacy of repair and/or progression of congenital or acquired heart disease. As in adult medicine, awareness and discussion of weight control, exercise, and other lifestyle issues must become an important part of the evaluation of obese children with heart disease in the cardiology outpatient setting. Weight may be one of the few modifiable risk factors for children with heart disease.

Pediatric cardiologists are frequently asked to provide recommendations for physical activity restriction in their patients with congenital and acquired heart disease or potentially life-threatening arrhythmias. These recommendations are extrapolated from published guidelines for competitive athletics in children with heart disease. In a retrospective study reviewing patients over a median of 8 years, Stefan et al reported that activity restriction in children with congenital heart disease was associated with the development of obesity. Even children who were of healthy weight at baseline had a higher risk of becoming obese over time if their activity was restricted. Physical activity limitation is a risk factor unique to children with heart disease. Importantly, physical activity restrictions in children with heart disease are not solely determined by practitioner recommendations. Indeed, these limitations may sometimes be initiated by parents or be self-imposed. Children with heart disease are often sedentary even when not limited by their physiology. Massin et al reported recently that children who had undergone the arterial switch operation were much less likely than their peers to participate in moderate or vigorous activity even when no restrictions had been placed by their cardiologists. Decreased activity may lead to deconditioning, decreased exercise capacity, and lower quality of life. A sedentary lifestyle associated with congenital heart disease is known to carry into adulthood and predict increased morbidity and mortality in this population.

Several recent studies have shown the benefits of physical training programs in both adults and children with congenital heart disease. Practitioners may need to refocus counseling during outpatient visits, providing careful instructions for appropriate and safe exercise regimens with regard to the underlying condition, in addition to more traditional counseling regarding exercise restrictions. Given the fact that patients with acquired and congenital heart disease have not escaped the epidemic of obesity, it is especially important for practitioners to adapt current activity guidelines from the American Heart Association and Bethesda conference for this purpose.

STUDY LIMITATIONS

Although the overall study population was quite large, small numbers in the transplant subgroup resulted in limited power to compare this group with the control subjects. Data collection was limited to children’s hospitals in Boston and Philadelphia and may not account for regional differences in obesity rates in other areas in the United States. In fact, control patients from CHB had slightly lower prevalence of overweight and obesity than those from CHOP, suggesting a slight difference in the underlying rate of pediatric obesity between the 2 cities. The data are also subject to the referral biases of each center. Data on race and its impact on the prevalence of obesity were not collected as part of this study.

It should be emphasized that data were limited to the information contained in the dictated clinic letters and notes. Therefore, we were unable to assess whether obesity and weight counseling may have occurred during the office visit without documentation. Important contributing data, such as parental weight and patient
activity levels, were not collected in this retrospective study, because this information was not consistently found in the clinic letters. In addition, we were only able to evaluate 1 comorbidity of obesity, blood pressure, from the data collected. In this population, the impact of the underlying heart disease on blood pressure may have been a confounder to its relationship to weight. All of the measurements used for analysis were obtained at only 1 point in time. This is especially important with regard to potential variability associated with blood pressure measurements in an outpatient office setting. Medications that might affect the outcome variables (eg, antihypertensive medication and steroids) were also not assessed in this analysis.

CONCLUSIONS
Obesity is a common significant additional risk factor for long-term cardiovascular disability in children with congenital and acquired heart disease, a population that is already at increased risk of shortened life expectancy. Although this study has important implications for the practice of outpatient pediatric cardiology, additional investigation is required to identify both common and potentially unique contributors to this problem, such as formal or perceived exercise restriction. Additional study is also required to truly understand the impact of overweight on the long-term outcome and cardiovascular health of children with underlying heart disease. Appropriate interventions for obese pediatric cardiac patients need to be developed. Given the known benefits of normal weight and exercise participation, advising against inactivity, obesity, and other unhealthy lifestyle choices and communicating these concerns to the referring physician should be important parts of the pediatric cardiologist’s care of children with cardiac disease.

REFERENCES
24. Hauser M, Bengel FM, Kuhn A, et al. Myocardial perfusion and coronary flow reserve assessed by positron emission tomogra-
ph in patients after Fontan-like operations. Pediatr Cardiol. 2003;24:386–392


35. Wray J, Sensky T. How does the intervention of cardiac surgery affect the self-perception of children with congenital heart disease? Child Care Health Dev. 1998;24:57–72


Effect of Vaccine Shortages on Timeliness of Pneumococcal Conjugate Vaccination: Results From the 2001–2005 National Immunization Survey

Philip J. Smith, PhD*, J. Pekka Nuorti, MD, DSc*, James A. Singleton, MS*, Zhen Zhao, PhD*, Kirk M. Wolter, PhD*

*National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia; National Opinion Research Center and University of Chicago, Chicago, Illinois

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

BACKGROUND. In September 2001 and again in February 2004, the Centers for Disease Control and Prevention announced shortages in the supply of the pneumococcal conjugate vaccine. We describe the effects of the pneumococcal conjugate vaccine shortages in 2001–2003 and 2004 on the timeliness of vaccination uptake for quarterly birth cohorts affected by the shortages.

METHODS. A total of 102,478 19- to 35-month-old children were sampled by the National Immunization Survey between 2001 and 2005. Provider-reported vaccination histories were used to evaluate whether children had been administered ≥4 doses of pneumococcal conjugate vaccine by 16 months of age.

RESULTS. Among successive birth cohorts affected by the first shortage, estimated coverage of ≥4 doses of pneumococcal conjugate vaccine by 16 months declined significantly from 28.8% to 18.2%. As the first shortage ended, estimated coverage of ≥4 doses of pneumococcal conjugate vaccine by 16 months increased steadily with each successive birth cohort to 40.2%. From the onset of the second shortage, estimated coverage of ≥4 doses of pneumococcal conjugate vaccine by 16 months declined steadily and significantly to 13.7%. As many as 27% of parents whose child was affected by the first shortage reported that their child’s vaccination provider had delayed the administration of pneumococcal conjugate vaccine doses. Of those parents who said that a pneumococcal conjugate vaccine dose was delayed and whose child was not administered ≥4 doses, 2.9% received a reminder notice from the provider to schedule administration of those delayed doses, and 0.2% had an appointment to receive those delayed or missed doses.

CONCLUSIONS. Vaccine shortages can result in delayed or missed doses and can have a dramatic impact on the vaccine coverage of children. Vaccination providers need to communicate effectively with parents so that doses that are delayed or missed during a vaccine shortage are administered when the shortage is resolved.
**Streptococcus pneumoniae** is a leading cause of bacteremia, sepsis, meningitis, pneumonia, sinusitis, and acute otitis media in young children. In February 2000 a 7-valent pneumococcal conjugate vaccine (PCV7), Prevnar (Wyeth-Ayerst Pharmaceuticals, Marietta, PA), was licensed. A total of 4 doses of PCV7 at ages 2, 4, 6, and 12 to 15 months was subsequently recommended for routine administration young children by the American Academy of Pediatrics in August 2000 and the Advisory Committee on Immunization Practices (ACIP) in October 2000. The routinely recommended PCV7 4-dose series has been 97% effective (95% confidence interval [CI]: 76%–100%) against invasive disease caused by serotypes represented in the vaccine, and effectiveness in children who received 3 doses before 12 months of age has been 87% (95% CI: 71%–94%).

In August 2001, deliveries of the vaccine were delayed, resulting in the first shortage of PCV7, and in December 2001, the ACIP recommended deferring the fourth dose of the vaccine for health children. In May 2003, the Centers for Disease Control and Prevention (CDC) reported that PCV7 production and deliveries had become sufficiently adequate to permit a return to the 4-dose schedule.

However, in December 2003, the CDC reported that it had received notice from the only US supplier of PCV7 that production constraints could cause delays in shipments. Subsequently, in February 2004, the CDC recommended that health care providers temporarily suspend routine use of the fourth dose of PCV7 to conserve the vaccine and minimize the likelihood of shortages, and then in March 2004, the CDC recommended that all health care providers temporarily suspend routine administration of both the third and fourth doses of PCV7 to healthy children. In September 2004, the CDC announced the end of the second PCV7 shortage because of the increased production capacity of the US supplier of PCV7 and reinstated the original American Academy of Pediatrics and ACIP recommendation of a 4-dose schedule to be completed by 16 months of age.

Little is known about the impact of the shortages on the timeliness of PCV7 vaccinations. Because the routine schedule recommends 3 doses of PCV7 by 7 months of age and 4 doses by 16 months, we evaluated the impact of vaccine shortages on the uptake of ≥3 doses of PCV7 by 7 months and ≥4 doses of PCV7 by 16 months of age. Also, we present estimated coverage of ≥3 and ≥4 doses of PCV4 by 24 months to evaluate subsequent uptake after those recommended ages. We also identified the characteristics of children whose PCV7 coverage was significantly lower as a result of the shortages. Finally, we examined the extent to which the parents of children affected by PCV7 shortages had reported delayed PCV7 doses, had an appointment to receive missed doses, or had ever received reminder notices to get missed doses.

**METHODS**

**Design of the National Immunization Survey**

Data collected from 102 478 19- to 35-month-old children sampled by the 2001, 2002, 2003, 2004, and 2005 National Immunization Survey (NIS) were analyzed. The NIS is a survey of US children 19 to 35 months of age conducted by the CDC for the purposes of monitoring vaccination coverage rates in the United States. Data collection in the NIS occurs in 2 phases, including a list-assisted random-digit dialing (RDD) survey of households with land-line telephones that have children 19 to 35 months of age, followed by a vaccination provider record check. Cellular telephone numbers are not included in the list of telephone numbers from which the RDD sample is obtained. When a household with an age-eligible child is identified in the RDD phase of the survey, an RDD interview is conducted that collects demographic information about each age-eligible child in the household, demographic information about the age-eligible child’s mother, and sociodemographic information about the household. At the end of the RDD interview, consent is requested for contacting the age-eligible children’s vaccination providers. If consent is given, the provider record-check phase of the NIS is conducted. In the provider record-check phase, all of the vaccination providers named by the RDD respondent are contacted by mail to obtain the age-eligible child’s provider-reported vaccination history. Provider-reported vaccination histories obtained from the provider record check were used to evaluate the vaccination status of children sampled in the NIS. The NIS was reviewed and approved by the institutional review board at the CDC in 2001 and 2006. Zell et al and Smith et al provide a detailed description of the statistical methods used by the NIS.

**NIS Coverage of the Population of Children 19 to 35 Months of Age and Survey Response Rates**

Between 2001 and 2005, among households with children 19 to 35 months of age, the percentage of 19- to 35-month-old children who lived in a household with a land-line telephone was ~90%. The response rate of the NIS is the product of 3 proportions: (1) the estimated percentage of households that reported in the NIS RDD interview as having a 19- to 35-month-old child among those that actually have a 19- to 35-month-old child, (2) the Council of American Survey Research Organizations rate of the RDD portion of the NIS, and (3) the percentage of sampled children for whom a provider-reported vaccination history is obtained in the NIS provider record check that is sufficiently detailed to accept as a complete report. The Council of American Survey Research Organizations rate is the multiplicative product of the resolution completion rate, the screener completion rate, and the RDD interview completion rate. The resolution completion rate is the proportion of telephone numbers determined to be residential,
nonresidential, or nonworking among all telephone numbers in the released NIS sample; the screener completion rate is the proportion of telephone number callers who completed the NIS screening interview among all of the resolved residential telephone numbers; and the RDD interview completion rate is the proportion of households that completed the NIS interview among all of the households with ≥1 child who was 19 to 35 months of age.

For the survey years that we studied, among households that had a land-line telephone and a 19- to 35-month-old child, the estimated percentages of households that reported having a 19- to 35-month-old child in the RDD portion of the NIS ranged from 74% and 70%, and the Council of American Survey Research Organizations rates ranged from 69% to 76%. Among parents of age-eligible children who had a completed NIS RDD interview, the percentages of children who had a sufficiently detailed vaccination history returned from vaccination providers to accept as a complete report ranged from 62% to 68%.

Statistical Analyses of the Timeliness of PCV7 Administration

Studies have demonstrated that the recommended PCV7 4-dose series administered before 16 months of age has very high effectiveness against invasive disease caused by serotypes represented in the vaccine.1 Although the ACIP specified alternative vaccination schedules for older infants and for catch-up doses during the initial uptake of the vaccine, we evaluated PCV7 uptake by the ACIP vaccination schedule for previously unvaccinated infants that recommends the routine administration of 4 doses of PCV7 by 16 months of age.1 During the PCV7 shortages, most of the children who were affected were between 2 and 15 months of age, the youngest and oldest ages at which doses are recommended in the routine 4-dose schedule. To measure the temporal effect of PCV7 shortages on the timeliness of PCV vaccination coverage among children affected by those shortages, we evaluated PCV7 coverage for quarterly birth cohorts. The quarterly birth cohort of each sampled child was defined by the year and quarter in which they were born. For each quarterly birth cohort, we estimated the percentage of children who were administered ≥4 doses of PCV7 by 16 and 24 months of age and the percentage of children who were administered ≥3 doses of PCV7 by 7 and 24 months of age.

To examine the extent to which parents of children affected by PCV7 shortage had reported a delayed PCV7 dose, had an appointment to receive missed doses, or had ever received reminder notices to get missed doses, we used data collected from the 2840 children sampled in the first vaccine shortage module of the NIS. Because the vaccine shortage module collected data between the second quarter of 2003 and the fourth quarter of 2003 from 19- to 35-month-old children, information collected in that module included data from quarterly birth cohorts who were to be administered the recommended PCV7 vaccination schedule during the first PCV7 shortage that was announced in August 2001 and lasted until May 2003.

In all of our analyses, we used SAS software survey procedures (SAS Institute, Inc, Cary, NC)14 that allow the sampling weights, sampling design of the NIS, independence of sampling from year to year, and clustering within households to be taken into account in our statistical analyses. For each quarterly birth cohort, we used a cohort-age model described by Smith et al11 to estimate the percentages of children who had received ≥4 PCV7 doses by 16 and 24 months of age and ≥3 PCV7 doses by 7 and 24 months of age. The statistical significance of differences between estimated percentages was evaluated using z-score tests.

RESULTS

Decreases in ≥4-Dose PCV7 Coverage by 16 Months During the Vaccine Shortages

Children affected by the first PCV7 shortage belonged to quarterly birth cohorts that were born approximately between the fourth quarter of 2000 and the fourth quarter of 2001 (Fig 1). Subsequent to the peak in vaccination coverage that preceded the first shortage, ≥4-dose PCV7 coverage by 16 months of age declined significantly (P < .05) from 28.8% (95% CI: ±1.9%) for the quarterly birth cohort born in the third quarter of 2000 to 18.2% (95% CI: ±1.7%) for the quarterly birth cohort born in the second quarter of 2001. Also, children affected by the second PCV7 shortage belonged to quarterly birth cohorts that were born approximately between the fourth quarter of 2002 and the third quarter of 2003 (Fig 1). Subsequent to the second peak in vaccination coverage that preceded the second shortage, ≥4-dose PCV7 coverage declined significantly (P < .05) from 40.1% (95% CI: ±2.8%) for the birth cohort born in the fourth quarter of 2002 to 13.7% (95% CI: ±1.2%) for the birth cohort born in the second quarter of 2003.

Late Doses of ≥4 Doses of PCV7 Administered by 24 Months

For all of the quarterly birth cohorts, coverage of ≥4 doses of PCV7 increased appreciably between 16 and 24 months of age (Fig 1). This increase was ~13% during the first PCV7 shortage and ~34% during the second shortage. However, during the first shortage, uptake of ≥4 doses of PCV7 by 24 months of age remained flat at ~33%. However, during the second shortage, coverage of ≥4 doses of PCV7 by 24 months of age decreased significantly from 51.2% (95% CI: ±3.4%) to 36.4% (95% CI: ±3.0%; P < .05).

≥3 Doses of PCV7 Coverage by 7 and 24 months

Vaccination coverage of ≥3 doses of PCV7 by 7 months of age was affected by the 2 PCV7 shortages (Fig 2). Subsequent to the first peak in vaccination coverage preceding the first shortage, ≥3-dose PCV7 coverage by 7 months of age declined significantly (P < .05) from...
FIGURE 1
Estimated vaccination coverage of ≥4 doses of PCV7 by 16 (dashed line) and 24 (dotted line) months of age among selected birth cohorts. Sample sizes for the coverage estimates at 16 months are listed directly below the estimate, and sample sizes for the coverage estimates at 24 months are listed directly above the estimate. Data are from the 2001–2005 NIS.

FIGURE 2
Estimated vaccination coverage of ≥3 doses of PCV7 by 7 (dashed line) and 24 (dotted line) months of age among selected birth cohorts. Sample sizes for the coverage estimates at 7 months are listed directly below the estimate, and sample sizes for the coverage estimates at 24 months are listed directly above the estimate. Data are from the 2001–2005 NIS.
43.2% (95% CI: ±2.2%) for the quarterly birth cohort born in the first quarter of 2001 to 27.7% (95% CI: ±2.1%) for the quarterly birth cohort born in the second quarter of 2002. Subsequent to the second peak in vaccination coverage preceding the second shortage, ≥3-dose PCV7 coverage declined significantly (P < .05) from 56.9% (95% CI: ±2.9) for the birth cohort born in the second quarter of 2003 to 15.9% (95% CI: ±2.8%) for the birth cohort born in the fourth quarter of 2003.

For all of the quarterly birth cohorts, vaccination coverage of ≥3 doses of PCV7 increased appreciably between 7 and 24 months of age (Fig 2). During the first PCV7 shortage, this increase was ~30% for birth cohorts. However, during the second PCV7 shortage, this increase was as much as 65%.

Association Between ≥4 Doses of PCV7 by 16 Months of Age and Characteristics of Children Affected by the First PCV7 Shortage

Among children affected by the first PCV7 shortage, the quarterly birth cohort born in the second quarter of 2001 had the lowest percentage of children administered ≥4 doses of PCV7 by 16 months of age. Among children in that birth cohort, Hispanic, non-Hispanic black, and non-Hispanic American Indian children had significantly lower vaccination coverage than non-Hispanic white children; foreign-born children had significantly lower vaccination coverage than native-born children; and children who had received their vaccine doses from all public, hospital, military, or mixed-type providers had significantly lower vaccination coverage than children who received all of their doses from a private vaccination provider (Table 1).

Also, children whose mother was never married or widowed/divorced/separated had significantly lower vaccination coverage than children whose mother was married; children whose mother had ≦12 years of schooling or had not completed college had significantly lower vaccination coverage than children whose mother was a college graduate; children whose mother preferred to speak Spanish during the NIS telephone interview had significantly lower vaccination coverage than those whose mother spoke English; and children whose mother was ≦29 years of age had a significantly lower vaccination coverage than those whose mother was ≦30 years of age.

Finally, children living in a household with annual income that was <135% of the federal poverty level had significantly lower vaccination coverage than children living in a household with an annual income above the 135% the federal poverty level; children who lived in a household with ≥2 children ≦18 years of age had significantly lower vaccination coverage than children with 1 child; and children who lived in a noncentral city metropolitan statistical area (MSA) or in a non-MSA had significantly lower vaccination coverage than children who lived in a central city MSA.

With respect to late doses received by 24 months of age, nearly all of the child, maternal, and household characteristics found to be statistically significant in explaining the variation in the percentage of children receiving ≥4 doses of PCV7 by 16 months were significant in explaining the variation in the percentage of children receiving ≥4 doses of PCV7 by 24 months (Table 1).

Association Between Coverage of ≥4 Doses of PCV7 by 16 Months of Age and Characteristics of Children Affected by the Second PCV7 Shortage

Among children affected by the second PCV7 shortage, the quarterly birth cohort born in the second quarter of 2003 had the lowest percentage of children administered ≥4 doses of PCV7 by 16 months of age. Compared with the quarterly birth cohort born in the second quarter of 2001, for the quarterly birth cohort born in the second quarter of 2003, there were relatively few factors associated with variation in vaccination coverage with ≥4 doses of PCV7 by 16 months of age. In particular, children’s race/ethnicity was not significantly associated with variation in estimated vaccination coverage of ≥4 doses of PCV7 by 16 months of age (Table 1).

However, foreign-born children had significantly lower vaccination coverage than native born children; and children who were administered all of their PCV7 doses at a public facility had significantly lower coverage than children who were administered all of their PCV7 doses at a private facility. Also, children whose mother had >12 years of education but not a college degree had significantly lower estimated coverage than children whose mother had a college degree. Finally, among children born in the second quarter of 2003, there were no household characteristics found to be significantly associated with variation in ≥4-dose PCV7 coverage by 16 months of age.

However, with respect to receiving ≥4 doses of PCV7 by 24 months, for children belonging to the quarterly birth cohort born in the second quarter of 2003, nearly all of the child, maternal, and household characteristics found to be statistically significant in explaining the variation in the percentage of children receiving ≥4 doses of PCV7 by 16 months of age among children born in the second quarter of 2001 reappeared as being statistically significant (Table 1).

Reminder/Recall Efforts to Administer Delayed and Missed Doses

Compared with the birth cohort born in the second quarter of 2000 that was scheduled to receive all of the recommended doses of PCV7 before the first shortage, birth cohorts that were scheduled to receive doses of PCV7 during the first shortage had significantly higher percentages of parents reporting that PCV7 doses had been delayed because the child’s vaccination provider told them that he/she was out of the vaccine (P < .05; Table 2). Among parents who reported a delay in administration of a PCV7 dose to their child and whose
During the First and Second PCV7 Shortages According to Selected Child, Maternal, and Household Characteristics: NIS 2001–2005

### TABLE 1

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>First PCV7 Shortage: Second Quarter of 2001</th>
<th>Second PCV7 Shortage: Second Quarter of 2003</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimated Percentage of Children Who Received ≥4 Doses of PCV7 by 16 mo, % (95% CI)</td>
<td>Estimated Percentage of Children Who Received ≥4 Doses of PCV7 by 24 mo, % (95% CI)</td>
</tr>
<tr>
<td><strong>Children's characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity of the Children</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>14.4 (± 3.8)*</td>
<td>12.0 (± 4.6)</td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>21.4 (± 2.3)*</td>
<td>14.8 (± 2.8)*</td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>13.0 (± 4.2)*</td>
<td>11.3 (± 5.2)</td>
</tr>
<tr>
<td>Non-Hispanic American Indian</td>
<td>9.3 (± 8.1)*</td>
<td>13.3 (± 15.4)</td>
</tr>
<tr>
<td>Asian, non-Hispanic</td>
<td>18.7 (± 8.6)</td>
<td>18.3 (± 10.7)</td>
</tr>
<tr>
<td><strong>Gender of the children</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>18.5 (± 2.4)*</td>
<td>12.4 (± 2.7)*</td>
</tr>
<tr>
<td>Female</td>
<td>18.0 (± 2.5)</td>
<td>15.2 (± 3.3)</td>
</tr>
<tr>
<td><strong>Birth country</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign born</td>
<td>9.1 (± 3.5)*</td>
<td>4.4 (± 8.7)*</td>
</tr>
<tr>
<td>Native born</td>
<td>14.7 (± 5.3)*</td>
<td>13.7 (± 2.1)*</td>
</tr>
<tr>
<td><strong>Vaccination provider type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All public facilities</td>
<td>9.1 (± 3.5)*</td>
<td>9.1 (± 4.7)*</td>
</tr>
<tr>
<td>All hospital facilities</td>
<td>14.7 (± 5.3)*</td>
<td>9.9 (± 6.4)</td>
</tr>
<tr>
<td>All private facilities</td>
<td>21.8 (± 2.4)*</td>
<td>14.7 (± 2.9)*</td>
</tr>
<tr>
<td>All military/other facilities</td>
<td>8.9 (± 8.1)*</td>
<td>20.8 (± 26.3)</td>
</tr>
<tr>
<td>Mixed (≥1 of the above categories)</td>
<td>15.4 (± 5.2)*</td>
<td>15.4 (± 9.0)</td>
</tr>
<tr>
<td><strong>Mother's characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother's marital status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widowed/divorced/separated</td>
<td>12.6 (± 5.5)*</td>
<td>14.6 (± 7.2)</td>
</tr>
<tr>
<td>Never married</td>
<td>11.8 (± 3.5)*</td>
<td>13.1 (± 5.9)</td>
</tr>
<tr>
<td>Married</td>
<td>20.9 (± 2.1)*</td>
<td>13.9 (± 2.2)*</td>
</tr>
<tr>
<td><strong>Mother's educational attainment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;12 y</td>
<td>11.7 (± 4.0)*</td>
<td>11.7 (± 4.9)</td>
</tr>
<tr>
<td>12 y</td>
<td>13.5 (± 3.0)*</td>
<td>12.5 (± 4.5)</td>
</tr>
<tr>
<td>&gt;12 y, noncollege graduate</td>
<td>18.5 (± 4.0)*</td>
<td>9.6 (± 3.6)*</td>
</tr>
<tr>
<td>College graduate</td>
<td>26.3 (± 3.0)*</td>
<td>17.3 (± 2.9)</td>
</tr>
<tr>
<td><strong>Mother's preferred language</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>19.1 (± 1.8)*</td>
<td>13.5 (± 2.1)*</td>
</tr>
<tr>
<td>Spanish</td>
<td>13.4 (± 5.2)*</td>
<td>15.0 (± 8.3)</td>
</tr>
<tr>
<td><strong>Mothers’ age, y</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤19</td>
<td>7.6 (± 5.3)*</td>
<td>4.4 (± 4.0)*</td>
</tr>
<tr>
<td>20–29</td>
<td>14.8 (± 2.5)*</td>
<td>11.6 (± 3.5)</td>
</tr>
<tr>
<td>≥30</td>
<td>21.7 (± 2.5)*</td>
<td>15.7 (± 2.7)</td>
</tr>
<tr>
<td><strong>Household characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income/poverty ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤135% of the federal poverty level</td>
<td>14.9 (± 3.4)*</td>
<td>14.7 (± 5.2)</td>
</tr>
<tr>
<td>&gt;135% of the federal poverty level</td>
<td>20.7 (± 2.1)*</td>
<td>13.6 (± 2.1)*</td>
</tr>
<tr>
<td>No. of children in the household ≤18 y of age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 child</td>
<td>24.9 (± 3.8)*</td>
<td>36.4 (± 4.2)*</td>
</tr>
<tr>
<td>2–3 children</td>
<td>17.9 (± 2.3)*</td>
<td>30.9 (± 2.7)*</td>
</tr>
<tr>
<td>≥4 children</td>
<td>15.6 (± 4.2)*</td>
<td>25.5 (± 5.9)*</td>
</tr>
<tr>
<td><strong>MSA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSA, central city</td>
<td>22.2 (± 3.5)*</td>
<td>12.2 (± 2.6)*</td>
</tr>
<tr>
<td>MSA, noncentral city</td>
<td>17.5 (± 2.1)*</td>
<td>17.1 (± 4.1)*</td>
</tr>
<tr>
<td>Non-MSA</td>
<td>14.9 (± 5.0)*</td>
<td>8.8 (± 3.0)*</td>
</tr>
</tbody>
</table>

*The estimated coverage rate for this level is significantly different from the estimated coverage rate of the reference category, denoted by an “*”.

** Nineteen- to 35-month-old children living in a household with an annual income ≤135% of the federal poverty level may be eligible for Medicaid if other state-specified criteria are also met.

---

child had not received ≥4 doses of PCV7 by the time of NIS RDD interview, only 2.9% (95% CI: ±2.6%) reported receiving a postcard, letter, or telephone call to remind them to bring their child back to the vaccination provider for this vaccine dose that was delayed, and 0.2% (95% CI: ±0.5%) reported having an appoint-
ment to have the missed PCV7 doses administered to their child.

**DISCUSSION**

In evaluating the impact of the 2 PCV7 shortages, we have shown that after the peaks in estimated vaccination coverage that preceded the 2 PCV7 vaccine shortages, coverage of ≥4 doses of PCV7 by 16 months of age declined steeply by 10.6% (95% CI: ±2.6%) in the first shortage and then by 26.4% (95% CI: ±3.0%) among quarterly birth cohorts affected by the second shortage. Our findings show that there was a large number of late doses administered to quarterly birth cohorts who were affected by the PCV7 shortages. However, we also found that a large percentage of children affected by the shortages had not received ≥4 doses of PCV7 by 24 months. Despite the shortages and associated lower coverage and delay of recommended doses, incidence of invasive pneumococcal disease has declined in young children after the introduction of PCV7. The additional disease burden that could have been prevented in the absence of the shortages is unknown, particularly among birth cohorts that were affected most by the shortage and among children who never received catch-up doses.

However, we also found disparities in estimated ≥4-dose PCV7 coverage between non-Hispanic white children and Hispanic, non-Hispanic black, and non-Hispanic American Indian children. In particular, the disparities we observed are similar to those observed in other reports on vaccination coverage surveillance and are often attributable to barriers to accessing primary care, such as not having health insurance, not having a medical home, or living in a location that is inconvenient to public or private clinics where vaccines are routinely administered. These disparities were found in the estimated vaccination coverage of ≥4 doses of PCV7 by 16 and 24 months during the first shortage and by 24 months in the second shortage. Although we did not find a disparity in the estimated coverage of ≥4 doses by 16 months between racial/ethnic groups in the second shortage, children who received all of the doses at a public facility had significantly lower estimated coverage than children who received all of the doses at a private facility.

In other research, a survey of 1412 vaccination providers found that 64% reported having a system to track children for whom PCV7 doses were deferred because of the shortages or because of the reduced number of vaccine doses recommended during the PCV7 shortages. However, our findings showed that, among children who had not received 4 doses of PCV7 and whose parent reported a delay in administration of a PCV7 dose, <1% reported that they had been contacted by their child’s vaccination provider to have those missed doses administered. How might the discrepancy in the findings between the vaccination provider survey and the findings from the NIS be reconciled? Perhaps some provider’s systems were not used to track children for whom PCV7 doses were deferred because of the shortages, or perhaps parents did not remember receiving the message or the messages received by parents from providers to recall children who missed doses.

The NIS response rate suggests one potential limitation of our work: because of the potential for selection bias resulting from multiple levels of nonresponse to the NIS, our statistical estimates of vaccination coverage may not accurately represent that of the entire target population of children 19 to 35 months of age in the United States. Although weighting adjustments were used to mitigate the potential for selection bias resulting from nonresponse, the effectiveness of those adjustments is unknown. If response to the NIS is associated with higher vaccination coverage or greater access to vaccinations during vaccine shortages, our estimates may underestimate the impact of the PCV7 vaccine shortage to the extent that the weighting adjustments do not adequately account for those factors. However, research by Bartlett et al has shown that estimates of vaccination coverage obtained from the National Health Interview Survey are nearly identical to those obtained from the NIS. Because the National Health Interview Survey is a door-to-door survey with interviews that are administered face to face, there is no bias in their estimates that can be attributed to households not having land-line telephones. Because the NIS estimates have been nearly identical to the estimates obtained from the National Health Interview Survey, there is empirical evidence that the statistical adjustments used to account for that potential source of selection bias work adequately in the NIS.

**TABLE 2**

<table>
<thead>
<tr>
<th>Birth Cohort, Birth Year–Quarter</th>
<th>Percentage Reporting PCV7 Dose Ever Delayed, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000–2</td>
<td>5.1 (±5.4)</td>
</tr>
<tr>
<td>2000–3</td>
<td>13.2 (±6.8)</td>
</tr>
<tr>
<td>2000–4</td>
<td>15.2 (±6.6)</td>
</tr>
<tr>
<td>2001–1</td>
<td>21.2 (±6.9)</td>
</tr>
<tr>
<td>2001–2</td>
<td>24.5 (±7.0)</td>
</tr>
<tr>
<td>2001–3</td>
<td>32.8 (±10.7)</td>
</tr>
<tr>
<td>2001–4</td>
<td>22.2 (±6.5)</td>
</tr>
<tr>
<td>2002–1</td>
<td>19.1 (±7.0)</td>
</tr>
<tr>
<td>2002–2</td>
<td>14.9 (±6.5)</td>
</tr>
</tbody>
</table>
Another limitation of our work is that, although the ACIP specified alternative vaccination schedules for older infants and for catch-up doses during the initial uptake of the vaccine, in our work we evaluated PCV7 uptake by the ACIP vaccination schedule for previously unvaccinated infants that recommends the routine administration of 4 doses of PCV7 by 16 months of age and did not account for the possibility that some vaccination providers followed the alternative schedules that required fewer doses to be administered. In analyses not presented in this article, we have accounted for this nuance and found that results did not change appreciably, however.

A steady and reliable supply of childhood vaccines is critical to combating vaccine preventable diseases. Since 2000, there have been shortages of the tetanus and diphtheria booster vaccine\(^{26}\); the influenza vaccine\(^{27}\); the measles, mumps, and rubella vaccine\(^{28}\); the varicella vaccine\(^{29}\); PCV7\(^{4,7}\); and, more recently, the meningococcal conjugate vaccine.\(^{30}\) These shortages have demonstrated the vulnerability of the vaccine supply.

Manufacturing a vaccine is a complex, highly controlled process that can take several months to over a year, and, thus, increasing the number of doses that are available to be administered can take time. Because there is frequently just 1 manufacturer producing a particular vaccine, even short-term disruptions in a manufacturer’s production volume may create a shortage that may result in underimmunization. Rodewald et al\(^ {30} \) have described 4 measures that the CDC takes to improve vaccine supply. First, the demand for vaccine production is heavily influenced by recommendations of the Advisory Committee on Immunization Practices that develops written recommendations for the routine administration of vaccines to the pediatric and adult populations, along with schedules regarding the appropriate periodicity and dosage. Second, the CDC negotiates a federal contract that uses federal and state funds to purchase ~56% of the recommended childhood vaccines distributed in the United States. Third, the CDC administers grants to states to help in implementing and promoting immunization programs to improve vaccine coverage rates. Fourth, the CDC administers some stockpiles of vaccines to be used for disruptions in production and surges in demand.

The General Accounting Office has recommended coordinating efforts between government and industry to increase and maintain vaccine stockpiles that will be available to all children during shortages, monitoring childhood vaccine inventory in state Vaccine for Children depots that can be used by all children during periods of vaccine shortages, and monitoring the delivery of vaccines to stockpiles and adequacy of the amount of vaccines in those stockpiles.\(^ {31}\) Vaccine stockpiles can be used to interrupt disease outbreak situations and to ameliorate short-term production problems. In light of recent vaccine shortages and increased concerns about an influenza pandemic or bioterrorism event, expansion of the CDC’s stockpiles has become a pressing public health need.

In addition to strengthening vaccine stockpiles, we believe that effective communication between vaccine providers and parents during vaccine shortages about when missed vaccines can be caught up is crucial. It is known that vaccination providers are the single most important source of information where parents seek advice about vaccines.\(^ {32,33} \) Also, efforts to remind parents when pediatric vaccines are due or to recall parents when vaccines are overdue have been demonstrated to be effective in increasing vaccination coverage rates.\(^ {34} \) Guidelines communicating with parents about other barriers to vaccination, such as concerns about vaccine safety, have been published.\(^ {35} \) In our current era of recurrent vaccine shortages, it is desirable that additional guidelines be published for pediatricians that describe how to communicate effectively with parents about vaccine shortages and how to establish an effective plan with parents to make up missed or deferred doses for their children.

### CONCLUSIONS

Children who miss doses that are not caught up during or after periods of vaccine shortage represent a pool of children who are at increased risk of acquiring and transmitting vaccine-preventable diseases. Our findings suggest that, during the initial uptake of a newly introduced vaccine, a vaccine shortage can have a significant effect on the vaccination coverage of birth cohorts affected by the shortage. Furthermore, children belonging to minority race/ethnic groups and other children who have characteristics or live in circumstances that are traditionally associated with low vaccination coverage may be impacted more greatly by the shortage.

### REFERENCES

5. Centers for Disease Control and Prevention. Notice to readers: updated recommendations on the use of pneumococcal conju-


32. Gellin BG, Maibach EW, Marcuse EK. Do parents understand their children who have concerns about vaccine safety and vaccination? A national telephone survey. Pediatrics. 2006;118(5). Available at: www.pediatrics.org/cgi/content/full/118/5/e1287


Does Current Asthma Control Predict Future Health Care Use Among Black Preschool-aged Inner-City Children?

Hemant P. Sharma, MD, Elizabeth C. Matsui, MD, MHS, Peyton A. Eggleston, MD, Nadia N. Hansel, MD, MPH, Jean Curtin-Brosnan, MA, Gregory B. Diette, MD, MHS

ABSTRACT

OBJECTIVES. Factors predictive of future asthma must be identified among young inner-city children, who suffer disproportionately from asthma. We investigated whether current asthma control predicts future asthma-related health care use among inner-city preschool-aged children with asthma.

METHODS. A total of 150 inner-city preschool-aged children with asthma were followed prospectively for 6 months. At baseline, symptom frequency and reliever-medication use were assessed to classify children into National Asthma Education and Prevention Program–derived control categories. Long-term controller-medication use was also assessed, as well as asthma-related health care use at baseline and at 3 and 6 months.

RESULTS. The mean age was 4.4 years, 92% were black, and 39% reported long-term controller use. At baseline, 37% were classified as having mild-intermittent, 17% had mild-persistent, 21% had moderate-persistent, and 25% had severe-persistent asthma control. Significant changes in asthma control were observed over time, including 46% of children originally categorized with mild-intermittent asthma who had worsened asthma control by 3 months. Asthma control significantly predicted future health care use 3 months later but not 6 months later. Multivariable analyses showed that, once control status was known, reported use of long-term controller medication added little additional predictive value.

CONCLUSIONS. Among inner-city preschool-aged children, significant fluctuations in asthma control occur as early as 3 months after assessment. Poor control but not long-term controller-medication use is an independent predictor of future asthma-related health care use at 3 months but is not significantly predictive of 6-month outcomes. Therefore, clinicians caring for inner-city children with asthma should consider reassessing asthma control at least every 3 months to identify those at highest future risk and to provide early interventions.
In the United States, asthma is the most common chronic illness of childhood, causing significant levels of preventable morbidity, particularly among inner-city black populations.1 The toll of pediatric asthma is especially reflected in the high rates of asthma-related health care use (HCU) in this high-risk population. One key component of successful prevention of asthma-related HCU is the ability to accurately identify those children at greatest risk of future morbidity. Thereby, children at high-risk may be targeted, and treatment and follow-up plans can be tailored specifically to lower their future risk. However, not enough is known about which clinical factors predict the risk of future asthma-related HCU among children. For example, previous asthma-related HCU is a well-recognized predictor of future HCU among various populations, including children,2–6 but there are undoubtedly other important predictors to be identified, especially in the high-risk black pediatric population.

There is growing interest in the use of asthma control assessment as a predictor of future morbidity.7–11 Asthma control describes the degree to which current disease activity is minimized by treatment. Severity, meanwhile, refers to a patient’s inherent disease state, as reflected by the degree of symptoms before starting treatment. Several limitations to the concept of asthma severity exist, including the incomplete ability to describe the dynamic nature of asthma, the unclear value of the baseline assessment after treatment has begun, and lack of consideration of both frequency and intensity of asthma exacerbations.7 Classification by level of control, rather than severity, is, therefore, recommended in the most recent version of the Global Initiative for Asthma (GINA) guidelines.12 Although the development of validated measures of asthma control has progressed well in adults, such validated measures of control for young children are still anticipated. Asthma status among children is most often assessed by the National Asthma Education and Prevention Program (NAEPP) guidelines, which recommend evaluation of symptoms, bronchodilator use, and lung function to determine appropriate therapy.13 In practice, the severity ratings outlined in the NAEPP guidelines are often interpreted as evidence of current control. There is already evidence that classification by the NAEPP guidelines predicts current and future asthma morbidity in adults,14 but studies among children have only been cross-sectional and have not evaluated current control as a predictor of long-term future morbidity.15

Other asthma guidelines have encouraged consideration not only of current asthma symptoms, but also long-term controller (LTC) medication use.16 In these classification schemes, a patient with rare symptoms on high-dose inhaled corticosteroids is considered to have worse asthma than one with the same symptom frequency on low-dose inhaled corticosteroids. The importance of LTC use as an independent predictor of future asthma morbidity is unknown. In addition, it is unclear whether the additional consideration of LTC use in other guidelines enhances our predictive ability beyond symptom status alone.

The need to identify factors predictive of future asthma risk is especially urgent among inner-city black children, because these children suffer disproportionate asthma morbidity, and disparities have widened over the latter half of the 1990s.17 Therefore, the purpose of this study was to determine which clinical indicators predict future asthma-related HCU in a population of inner-city preschool-aged children with asthma.

METHODS

Study Population and Recruitment Procedures

Recruitment for this study, a part of the Center for Childhood Asthma in the Urban Environment, occurred between September 2001 and December 2003. Participants were recruited from the health systems that provide care to the majority of residents of east Baltimore. Inclusion criteria were (1) age between 2 and 6 years, (2) residence within the catchment area defined by 9 contiguous zip codes, (3) doctor-diagnosed asthma, and (4) symptoms of asthma and/or medication use for asthma in the previous 6 months.

During a screening visit, informed consent was obtained, and an extensive survey was conducted, followed 2 weeks later by administration of the clinical status questionnaire. Participants had 2 additional visits at 3 and 6 months for repeat assessments of asthma status and outcomes. The study was approved by the Johns Hopkins School of Medicine Institutional Review Board.

Clinical Status Questionnaire

During the baseline visit, a questionnaire was administered by trained study personnel to the primary caregiver to assess demographic characteristics, asthma-related HCU in the previous 3 months (unscheduled doctor [UD] visits, emergency department [ED] visits, and hospitalization), LTC use (within the previous 2 weeks), rescue medication use (within the previous 2 weeks), and recent symptoms (within the previous 2 weeks). Use of the following LTC medications was assessed: inhaled corticosteroids, cromolyn and nedocromil, oral leukotriene modifiers, long-acting β-agonists, and oral theophylline. HCU outcomes and LTC-medication use were analyzed as dichotomous variables. Ten close-ended questions assessed frequency of daytime and nighttime symptoms, short-acting β-agonist use, and activity limitation during the previous 2 weeks. Ordered responses to these questions, ranging from 0 to 14 (number of days in the last 2 weeks), were fit to the NAEPP scheme as closely as possible to assign participants to a given asthma control category.18 For example,
criteria for mild-intermittent control included wheezing, coughing, or chest tightness ≤2 days per week; mild-persistent control included >2 days per week but less than daily; and moderate-persistent control included daily. These symptoms represented control, rather than severity, because they reflected recent disease activity, irrespective of treatment. The follow-up questionnaires administered at 3 and 6 months assessed interval asthma-related health care use (previous 3-month period) and recent symptoms (within the previous 2 weeks).

**Statistical Analysis**

The relationships between asthma control at the baseline visit and recent asthma-related HCU were analyzed. In addition, relationships between asthma control at the baseline visit and future HCU, assessed at the 3- and 6-month visits, were analyzed. The $\chi^2$ test was used to compare proportions between groups, and logistic regression was used to generate odds ratios (ORs). The independent variables of interest were asthma control classification and LTC-medication use. Multivariate logistic regression was used to analyze the relationships between the independent variables and the 3 outcomes of interest: UD visit, ED visit, and hospitalization. In logistic regression models, asthma control was treated as a 4-category ordinal variable. Age, gender, and caregiver educational attainment were not found to be significant confounders and were, therefore, not included in the final models. All of the analyses were performed with Stata SE 8.0 (Stata Corp, College Station, TX). Statistical significance was defined as a $P$ value of <.05.

**RESULTS**

**Baseline Study Population**

The baseline cohort of 150 children had more boys (57%) than girls, was predominantly black (92%), and had caregivers mostly with highest educational attainment of high school or less (80%; Table 1). The mean age of the participants was 4.4 years. All 4 of the asthma control categories were well represented at baseline. Although 62% of the children met criteria for persistent asthma, only 39% reported using a LTC medication at baseline. Even among the subjects with severe persistent asthma, only 56% reported LTC-medication use. Of those reporting use of an LTC medication, 86% were taking inhaled corticosteroids, and the remaining 14% were taking oral leukotrienes or cromolyn/nedocromil. Recent outpatient asthma-related HCU was more common (UD visit: 18%; ED visit: 24%) than inpatient HCU (hospitalization 3%).

**Variability in Asthma Control Over Time**

Asthma control varied markedly over time (Fig 1). Nearly half (46%) of children originally categorized with mild-intermittent asthma at baseline had worsening in their degree of asthma control by 3 months, which would have required intensification of treatment according to guideline recommendations. Similarly, of those with mild-persistent asthma at baseline (Fig 1B), one third had worse asthma control by 3 months. Changes in the level of control between baseline and 3 months were also observed in more than half of those with moderate-persistent asthma (Fig 1C) and approximately half of those with severe persistent asthma (Fig 1D).

Even greater shifts were observed between baseline and 6 months among children with persistent symptoms at baseline. More than two thirds of children with persistent asthma experienced a change in their degree of asthma control by 6 months.

**Relationship of Current Asthma Control to Recent Asthma-Related HCU**

There was a strong, significant association between poor asthma control and recent HCU, suggesting that asthma control is related to overall recent disease activity. For example, the proportion of children with an asthma-related UD visit in the preceding 3 months increased significantly with worse control (mild-intermittent: 5%; mild-persistent: 8%; moderate-persistent: 23%; severe-persistent: 42%; $P < .01$ by Cuzick’s test for trend; Fig 2A) Similar trends were observed for asthma-related ER visits in the previous 3 months ($P < .01$) and asthma-related hospitalization, although the latter relationship

| TABLE 1 Baseline Characteristics of Children With Asthma (N = 150) |
|------------------------|------------------|
| Characteristic         | Data             |
| Age at initial visit, median (range), y | 4.4 (2.0–7.0) |
| Race/ethnicity, %      |                  |
| Black                  | 92               |
| White                  | 3                |
| Other                  | 5                |
| Gender, %              |                  |
| Male                   | 57               |
| Female                 | 43               |
| Educational attainment of caregiver, % |                  |
| Eighth grade/some high school | 39           |
| High school graduate   | 42               |
| Some college or more   | 19               |
| Asthma control category, % |        |
| Mild intermittent      | 38               |
| Mild persistent        | 18               |
| Moderate persistent    | 20               |
| Severe persistent      | 24               |
| Taking LTC medication (overall) by control status, % |                  |
| Mild intermittent      | 39               |
| Mild persistent        | 29               |
| Moderate persistent    | 27               |
| Severe persistent      | 47               |
| Recent HCU (previous 3 mo), % |            |
| UD visit               | 18               |
| ED visit               | 24               |
| Hospitalization        | 3                |
was not statistically significant ($P = .07$). In logistic regression models, there was approximately a twofold average increase in the odds of recent asthma-related HCU for each increment of worsening asthma control (Table 2).

Relationship of Current Asthma Control to Future Asthma-Related HCU
In addition to being associated with recent HCU, baseline asthma control also strongly predicted future asthma-related HCU at 3 months. (Fig 2B) Poorer baseline...
asthma control was significantly associated with an increasing likelihood of an interval 3-month UD visit (P < .01) or ED visit (P = .02). Similar trends were seen for 3-month hospitalization, although the relationship was not significant (P = .07). Despite being a good predictor at 3 months, asthma control did not predict future asthma-related HCU at 6 months, as shown: UD visit (P = .35), ED visit (P = .09), and hospitalization (P = .33; Fig 2C).

Because baseline control status was a good predictor of future asthma-related HCU during the subsequent 3 months but less so beyond this 3-month interval, additional analyses examined relationships between asthma control status at 3 months and HCU at 6 months. A key additional observation was that the poorer the asthma control at the 3-month visit, the greater the odds of asthma-related HCU at the 6-month visit. For example, there was a 1.7-fold increase in the odds of an ED visit at 6 months for each worse category of asthma control status when assessed at 3 months (OR: 1.7; 95% CI: 1.1–2.8), supporting the notion that control status predicts future HCU up to 3 months after assessment.

**Relationship of LTC-Medication Use to Recent and Future HCU**

Not surprisingly, baseline LTC-medications use was significantly associated with an approximately threefold average increase in the odds of recent HCU, specifically UD and ED visits (Table 2). However, baseline LTC use was a significant predictor of only 1 future outcome, UD visit at 3 months (OR: 5.5; 95% CI: 2.0–15.5).

**Multivariate Model of Asthma Control Status and Reported LTC-Medication Use**

After adjusting for baseline asthma control status, LTC-medications use had little additional predictive value for future HCU (Table 2). LTC use was only an independent predictor of UD visit at 3 months (OR: 4.0; 95% CI: 1.4–11.9). On the other hand, adjustment for LTC status had little impact on the associations between control status and future HCU. Independent of LTC use, baseline control predicted future HCU at 3 months. For example, relative to those with mild-intermittent baseline control, children with mild-persistent asthma had an approximately twofold increased odds of an ED visit by 3 months, children with moderate-persistent asthma had an approximately threefold increased odds, and children with severe-persistent asthma had an approximately fourfold increased odds (Fig 3). Neither baseline asthma control nor LTC use was associated with asthma-related HCU by 6 months.

**DISCUSSION**

This study found that current asthma control is a strong predictor of future asthma-related HCU among inner-city black children with asthma. In addition, asthma control status changed substantially in as little as 3 months, suggesting that asthma control should be reassessed frequently so that appropriate changes in treatment can be made with the aim of reducing future HCU.

Other studies have shown that asthma control predicts HCU in adult patients, but this is the first study, to our knowledge, that shows that baseline control predicts future HCU in one of the most vulnerable asthma populations, namely, inner city preschool minority children. Studies in adults have shown that poor asthma control is associated with both current and future ED visits, doctor’s visits, and hospitalizations. Although these studies in adult populations are informative, it is unclear whether these findings can be extrapolated to children, particularly inner-city black children who arguably are at highest risk for morbidity. Two cross-sectional studies in pediatric populations found relationships between asthma status and recent HCU, but future HCU was not examined. Another study found that control as assessed by diary-reported asthma symptoms predicted future HCU among children, but outcomes were assessed only in the short-term, 1 month into the future. Our study extends this body of litera-
ture in the pediatric population by focusing specifically on a high-risk group, black inner-city children, and by relating current guideline-derived asthma control status to longer-term future asthma-related HCU.

It is striking that current asthma control was found to be a strong predictor of future HCU in the subsequent 3 months but not beyond. The lack of longer-term predictive ability may be because of the inherently labile nature of asthma, particularly in the inner city population included in our study. Of particular concern is the fact that nearly half of the children initially classified with mild-intermittent asthma had worse control just 3 months later, crossing a threshold that would have required treatment escalation according to asthma guidelines. Therefore, reassessment, even as soon as 3 months after an initial evaluation, would have already missed an early opportunity for intervention. Taken together, these findings suggest that frequent reassessment of control status may reduce future morbidity in this high-risk population.

Other asthma guidelines, such as the 1995 GINA guidelines, have proposed examining not only symptom frequency but also LTC-medication use in the determination of asthma status. However, our study found that assessment of LTC-medication use, although essential for planning appropriate asthma management, offered little additional value to predicting future HCU once control status was already known. This finding is consistent with the most recent version of the GINA guidelines, which reflects a growing awareness that classification of asthma by level of control may be more useful in identifying high-risk patients than classification by underlying severity, which takes LTC use into account. Another question of interest would be whether the predictive ability of LTC use differs by type of LTC. However, given that almost all (86%) of the participants reporting LTC use were taking inhaled corticosteroids, we were unable to investigate that question.

It is notable that the most recent GINA guidelines recommend a dichotomized classification scheme for control (ie, poorly controlled versus controlled), which may be more simple to consider. However, our study’s findings suggest that there is value in considering asthma control status for levels of control, because the risk of future HCU, at least for the next 3 months, varies significantly from one control category to the next. Thus, oversimplification of control status may result in a loss of important clinical information.

Although LTC-medication use would be expected to be an indicator of underlying asthma severity and, consequently, risk of morbidity, this concept may not hold in a “real-world” setting. Outside of a controlled setting, such as a clinical trial, reported use of an LTC medication is likely a marker of overall quality of asthma care rather than disease severity. For example, reported LTC-medication use may reflect access to care, effective provider-patient communication, the health care provider’s clinical judgment, and ability to pay for the prescription, among many other factors. It is not surprising, then, that in our inner-city observational study, 62% of children had persistent asthma symptoms, but only 39% were on an LTC medicine. This finding highlights a strength of our study in its ability to look at actual use of LTC medication, a concept distinctly different from need for such medication. Reported LTC-medication use was only weakly related to clinical indications for such medication and, therefore, was a poor predictor of future HCU.

Recently, there has been significant attention focused on asthma control, rather than severity, assessment. This attention to control stems in part from limitations to the

**FIGURE 3**
Progressively worse baseline asthma control predicts significantly increased odds of an ED visit at 3 months, whereas baseline LTC-medication use is not a significant independent predictor of a 3-month ED visit. Also, the effect size of multistep changes in control is greater than that of LTC use. Adjusted ORs and 95% CIs were derived from multivariate modeling, including asthma control and reported LTC use. For asthma control, ORs refer to the odds of an ED visit in those of a given baseline asthma control class relative to those with asthma control at 1 step, 2 steps, and 3 steps improved, respectively.
operative concept of asthma severity and the underuse of severity guidelines in clinical practice. Several well-validated measures of asthma control have been developed, some of which have been found to predict future HCU among adults. These measures classify control based on symptom frequency, similar to our study, so that these other measures may also prove to be useful predictors of future HCU. In addition, other indicators of asthma control may predict future HCU, such as lung function, sputum eosinophilia, and exhaled nitric oxide. Future studies should focus on the joint assessment of symptom frequency, lung function, inflammatory markers, and other patient characteristics, which may enhance our ability to predict future asthma risk.

There are some limitations to this study to consider. All of the data regarding symptoms, medication use, and HCU were collected by caregiver report. Therefore, our findings regarding LTC use apply only to reported LTC use, which may differ from verified use, because medication compliance was not assessed. However, reported use may be of greater relevance, because provider determination of LTC use is often reliant on patient report. In addition, HCU as reported by caregivers may have been subject to limitations of recall. However, Juniper et al found questionnaires to be more reliable and responsive than daily symptom diaries, and the time frame of 3 months, used in our study for HCU recall, has been shown to be more precise than recall windows of longer duration. Finally, our findings may not be generalizable to other populations, such as nonblacks and those outside the urban setting. However, the findings have direct implications to inner-city black children, who bear much of the asthma burden in the United States.

CONCLUSIONS

We found that asthma control status but not reported LTC-medication use is a strong predictor of future asthma-related HCU among inner-city black preschool-aged children. Findings from this study underscore the labile nature of asthma, because marked fluctuations in the degree of asthma control were observed within only 3 months. Our findings emphasize that asthma control is valuable to assess in young inner-city children and suggest that such assessments should be repeated at least every 3 months to take advantage of opportunities to prevent future morbidity.

ACKNOWLEDGMENTS

This work was supported by US Environmental Protection Agency Science to Achieve Results program. Additional funding and support was provided by Johns Hopkins Center for Urban Environmental Health grant NIEHS P30 ES 03819.

REFERENCES

Predicting asthma morbidity in Harlem emergency department patients. *Acad Emerg Med.* 2004;11:944–950


A Cross-National Comparison of Racial and Ethnic Disparities in Low Birth Weight in the United States and England

Julien O. Teitler, PhD, Nancy E. Reichman, PhD, Lenna Nepomnyaschy, PhD, Melissa Martinson, MSW

School of Social Work, Columbia University, New York, New York; Department of Pediatrics, Robert Wood Johnson Medical School, New Brunswick, New Jersey

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. We used 2 new nationally representative surveys to compare racial and ethnic differences in low birth weight in the United States and England.

METHODS. Risk factors and rates of low birth weight were compared across groups for singleton births within each country (white, black, Hispanic, Asian, and American Indian mothers in the United States; white, black, and Asian mothers in England). Crude rates and rates adjusted for socioeconomic status and behaviors were compared. Additional comparisons were limited to native-born mothers.

RESULTS. Racial and ethnic disparities in low birth weight are as large in England as in the United States. Socioeconomic status and behaviors explain little of the variation across racial and ethnic groups in either country.

CONCLUSIONS. Health disadvantages associated with being a minority do not seem to be a uniquely American phenomenon. Universal health care, as provided in the United Kingdom, alone may be insufficient to reduce racial and ethnic disparities in low birth weight.
Health disparities are an important marker of inequality in a society. In the United States, there are large socioeconomic disparities in health. There are also large racial and ethnic health disparities, even among infants. In 2001, 13.0% of infants born to black mothers in the United States were low birth weight (<2500 g) compared with 6.7% of those born to white mothers, 6.5% of those born to Hispanic mothers (of any race), 7.3% of those born to American Indian mothers, and 7.4% of those born to Asian and Pacific Islander mothers. Although low birth weight is not a direct measure of infant morbidity, it is frequently used as a marker for poor health at birth because it is a leading risk factor for infant mortality and for subsequent morbidity among surviving infants. It is also accurately measured, widely available, and useful for sub-national and international comparisons.

Despite substantial research on determinants of racial and ethnic disparities in birth outcomes in the United States, much remains to be explained. Cigarette smoking during pregnancy is strongly associated with low birth weight but does not explain the black-white disparity. Factors that are related to low socioeconomic status (SES), such as young maternal age and low educational attainment, are associated with low birth weight but explain little of the black-white disparity.

The extent to which racial and ethnic disparities in low birth weight reflect socioeconomic inequalities remains an open question, because most representative data on births by race do not include detailed measures of SES and because the association does not seem to hold for certain groups. In the United States, Hispanic mothers as a broad group have favorable birth outcomes despite their low SES, and across virtually all racial and ethnic groups, foreign-born mothers have lower rates of low birth weight than their more advantaged US-born counterparts. Proposed explanations for the Hispanic and immigrant advantages involve selective migration on the basis of health and protective cultural factors.

Little is known about racial and ethnic disparities in birth outcomes in developed countries other than the United States, because very few countries collect vital statistics data by race and ethnicity. Only recently have national data on births by race become available in the United Kingdom, where a socioeconomic gradient in health similar to that in the United States exists despite substantial differences in the provision of health care and public assistance systems. The overall rate of low birth weight in the United Kingdom in 2000 was 7.6%, the same as the US rate in 2000. As in the United States, low birth weight in the United Kingdom is negatively associated with SES. A recent report indicated that racial and ethnic disparities in low birth weight do exist.

We used data from 2 new and remarkably comparable nationally representative birth cohort studies to compare disparities in low birth weight in the United States and England. We compared crude disparities and disparities adjusting for SES and prenatal behaviors. We were interested primarily in comparing the levels of within-country disparities. Cross-country comparisons of specific groups may be interesting, but it is critical to keep in mind that groups often referred to with the same label in the United States and England (eg, black or Asian) consist of individuals with very different histories and ancestries.

Background
Although they are both English-speaking Western developed countries with clear cultural and historical connections, the United States and United Kingdom have different public assistance systems, health care delivery systems, and racial and ethnic compositions. The United Kingdom, but not the United States, has a guaranteed minimum income and in 2001 spent 13.7% of its gross domestic product on “social protection spending” (including old age, survivor, disability, unemployment, welfare, and housing benefits); in contrast, 8% of gross domestic product was spent on social programs that same year in the United States. The United Kingdom has a tax-supported government entity, the National Health Service, that is responsible for providing comprehensive medical care to all residents, whereas the United States has both a private decentralized system and a set of publicly funded programs for specific segments of the population but does not guarantee universal health care.

In the United States, 75.1% of the population is white. The vast majority of black people in the United States, who represent 12.3% of the nation’s population, are descendents of African slaves. A minority (but an increasing number) of black people in the United States are immigrants from the Caribbean and sub-Saharan Africa. Asian people (primarily Chinese, Filipino, Indian, Vietnamese, Korean, and Japanese) comprise 3.6% of the US population; American Indian, 0.9%; other races, 5.6%; and ≥2 races, 2.4%. A large share (12.6% and growing) of the US population is Hispanic (of any race), mostly of Mexican descent. Finally, 19.9% of the US population is foreign born (5% of white population, 4.9% of black population, 2.3% of American Indian population, 63.1% of Asian population, and 35.8% of Hispanic population).

In the United Kingdom, 91% of the population is white. The majority of nonwhite people in the United Kingdom are black or Asian. Black people represent 2.3% of the population and are primarily of Caribbean or African ancestry. Asian people from the Indian subcontinent represent 4.6% of the United Kingdom population and are primarily Indian, Pakistani, and Bangladeshi. The remaining 2.2% of the population is of other races or ethnicities (which includes Chinese). In 2001, 8.3% of the United Kingdom population was born outside of the United Kingdom (a third of immigrants...
were born in other European countries). Because the bulk of Caribbean, Asian, and African immigration to the United Kingdom has occurred since World War II, most minority individuals in the United Kingdom are first- or second-generation immigrants.

Given the universal access to health care, minimum guaranteed income, and large representation of immigrants (who may be more favorably selected on the basis of health or behaviors than the native born) in minority groups in the United Kingdom, we expected that racial and ethnic disparities in low birth weight in England would not be as pronounced as those in the United States.

METHODS

Data

This article is based on data from 2 new nationally representative birth cohort studies, the Early Childhood Longitudinal Survey-Birth Cohort (ECLS-B) for the United States and the New Millennium Cohort Study (MCS) for England. The ECLS-B is a nationally representative sample of >10 000 children born in the United States in 2001. Births were sampled from birth certificates, and, for infants who were alive and residing in the United States at 9 months of age, the parents (usually mothers) were interviewed. The survey data were linked to the infants' birth records. The response rate was 74%. The sample is representative of 9-month-surviving infants born in the United States in 2001 to mothers ≥15 years old who did not place the child for adoption.

The US analyses are based on a subsample of 8300 singleton births with nonmissing data on all of the covariates. All unweighted ECLS-B sample sizes are rounded per contractual data user requirements.

The MCS is a birth cohort study of 18 553 children born in the United Kingdom in 2000–2001. It is representative of all children born in the United Kingdom between 2000 and 2001 who were living in the United Kingdom at 9 months of age and who were eligible to receive the Child Benefit (a universal benefit for all residents of the United Kingdom) at that time. A follow-up survey, which was conducted between September 2003 and April 2005, had an 80% response rate (of mothers who completed initial interviews). We limited our analyses to England because the ethnic minority sample sizes in Northern Ireland, Scotland, and Wales are very small.

We used a subsample of 10 599 singleton births in England with nonmissing data on all of the covariates. The ECLS-B and MCS are highly comparable, because they are from the same year, they are both nationally representative of 9-month infant survivors, and they both include data on many pertinent risk factors.

Measures

The outcome measure is low birth weight (<2500 g), which in the ECLS-B comes from the infants’ birth certificates and in the MCS comes from maternal reports of their infant’s birth weight. In the United States, maternal reports of birth weight are highly consistent with those from birth certificates.

Racial and ethnic categories that are meaningful or routinely used in each country were used. For the United States, the analysis groups were white, black, Hispanic, Asian, and American Indian. Hispanic mothers were not included in the other racial and ethnic categories. For England, the analysis groups were white, black (mostly of Caribbean and African descent), and Asian (from the Indian subcontinent (Indian, Pakistani, and Bangladeshi). The race and ethnicity data came from the main respondent in both surveys, and the birth mother’s race or ethnicity was used.

The following analysis variables are comparable in the 2 data sets: maternal age (<20 years, 20–34 years, and ≥35 years), parity (first birth versus higher-order birth), marital status, nativity (native versus foreign born), any prenatal care received, and any cigarette smoking during pregnancy. Both data sets also include information on alcohol consumption during pregnancy, but that behavior is not included in our analyses because of a high rate of missing data in the ECLS-B. The mother’s birthplace comes from the birth certificates for the ECLS-B (island-born Puerto Rican mothers are categorized as native born) and from the follow-up survey for the MCS. Prenatal care and smoking are from self-reports in the MCS and from birth certificates in the ECLS-B. For England, the measure of employment is any work activity during pregnancy, and for the United States, it is any work activity during the 12 months before the birth.

Measures of poverty and education that are meaningful for each country, but not directly comparable, were used. For the United States, if a family’s household income fell below the official poverty income threshold (adjusted by household size) for the survey year, that mother was considered poor. For England, the measure of poverty was receipt of means-tested benefits, because there was a high rate of missing data on income in the MCS.

The US education variable consisted of 4 categories: less than high school, high school diploma or General Education Development, some college, and a bachelor’s degree or higher. The United Kingdom education variable also consisted of 4 categories: less than O level (0 to 11 years of education), at least O level but less than A level (12–13 years of education), A level or higher, and other qualifications (eg, education obtained outside of the United Kingdom). As indicated above, these categories were not directly comparable with the US educational classifications.

Analyses

We compared the following across racial and ethnic groups in each country: (1) risk factors for low birth weight; (2) rates of low birth weight (crude, as well as adjusted for infant gender, SES, and behaviors); and (3) rates of low
birth weight among native-born mothers. All of the analyses were based on weighted data. The analyses of native-born mothers in the United States were based on the 5900 mothers in the ECLS-B analysis sample who reported in their baseline interview that they were born in the United States. The analyses of native-born mothers in England were based on the 7212 mothers in the MCS analysis sample who completed the follow-up survey and reported at that time that they were born in the United Kingdom. We assessed sensitivity of the findings with an alternative birth weight threshold and with an alternative poverty measure for the United States.

RESULTS

Racial and Ethnic Differences in Risk Factors in the United States and England

There were large differences in risk factors both within and across countries (Table 1). In the United States, non-Hispanic black, Hispanic, and American Indian mothers were more likely, and Asian mothers were less likely, than non-Hispanic white mothers to have births at young ages. In England, teen birth rates were considerably lower than those in the United States; the rate for white mothers was similar to that for black mothers and higher than that for Asian mothers.

In the United States, non-Hispanic white mothers were more likely than non-Hispanic black, Hispanic, or American Indian mothers, and less likely than Asian mothers, to be married when they gave birth. In England, black mothers were much less likely, and Asian mothers were much more likely, than white mothers to be married at the time of the birth.

In the United States, most Hispanic and Asian mothers were foreign born, and the vast majority of non-Hispanic white, non-Hispanic black, and American Indian mothers were native born. In England, <5% of white mothers were foreign born, whereas substantial proportions of Asian (61%) and black (43%) mothers were born outside of the United Kingdom.

In both countries, there were large differences in poverty status across racial and ethnic groups. In particular, black mothers were ~3 times as likely as white mothers to be classified as poor in both countries based on the country-specific measures that we used. In the United States, Hispanic and American Indian mothers were ~3 times as likely as non-Hispanic white mothers to be poor. The poverty rates of Asian mothers were fairly similar to those of white mothers in each country. The racial and ethnic patterns in educational attainment did not follow the patterns in poverty. In the United States, Hispanic mothers had lower rates of high school completion than black mothers, despite a lower poverty rate. In England, black mothers were as likely as white mothers to have an A-level or higher degree despite a

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>United States</th>
<th>England</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>.111</td>
<td>.069</td>
</tr>
<tr>
<td>20–34</td>
<td>.752</td>
<td>.757</td>
</tr>
<tr>
<td>≥35</td>
<td>.137</td>
<td>.174</td>
</tr>
<tr>
<td>First birth</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.410</td>
<td>.425</td>
</tr>
<tr>
<td>Married</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.676</td>
<td>.621</td>
</tr>
<tr>
<td>Foreign born</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.211</td>
<td>.099</td>
</tr>
<tr>
<td>Socioeconomic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.226</td>
<td>.173</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>.272</td>
<td>.251</td>
</tr>
<tr>
<td>High school/GED</td>
<td>.218</td>
<td>.332</td>
</tr>
<tr>
<td>Some college</td>
<td>.266</td>
<td>.375</td>
</tr>
<tr>
<td>Bachelors degree or more</td>
<td>.244</td>
<td>.021</td>
</tr>
<tr>
<td>Education (United Kingdom)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than O level</td>
<td>—</td>
<td>.251</td>
</tr>
<tr>
<td>O level to less than A level</td>
<td>—</td>
<td>.353</td>
</tr>
<tr>
<td>A level or higher</td>
<td>—</td>
<td>.375</td>
</tr>
<tr>
<td>Other qualification</td>
<td>—</td>
<td>.021</td>
</tr>
<tr>
<td>Behavioral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received prenatal care</td>
<td>.991</td>
<td>.975</td>
</tr>
<tr>
<td>Smoked cigarettes</td>
<td>.137</td>
<td>.225</td>
</tr>
<tr>
<td>Worked during pregnancy</td>
<td>.717</td>
<td>.685</td>
</tr>
</tbody>
</table>

Figures (except those listed in the final row) are proportions that are weighted and adjusted for design effects. — indicates no data; GED, General Education Diploma.
substantially higher rate of poverty, and Asian mothers had lower levels of educational attainment than white mothers despite similar poverty levels.

Prenatal care use was widespread in both countries across all of the groups. More than 90% of each racial and ethnic group in each country received some prenatal care. Other prenatal behaviors varied across racial and ethnic groups in both countries, but in a paradoxical way, because white mothers (who had the lowest rates of low birth weight) were more likely to report engaging in behaviors that are associated with low birth weight. In particular, cigarette smoking was high among white mothers in both countries. Asian mothers in both countries reported very low rates of smoking during pregnancy (3% and 5.5%, respectively, in the United States and England).

In sum, in both countries, minority mothers differed from each other and from white mothers in terms of risk factors for low birth weight. In the United States, Asian mothers shared the socioeconomic advantages of non-Hispanic white mothers and also had a low rate of cigarette smoking (ie, they had the most favorable risk factor profile of all of the groups). Black and American Indian mothers were at the greatest risk for low birth weight based on their risk factors, with Hispanic mothers falling between black and white mothers. American Indian mothers were the only group with a higher rate of cigarette smoking than white mothers. In England, Asian mothers had the most favorable risk factor profile of all of the groups, and black mothers, although more likely than white mothers to be poor, had other offsetting factors. They were less likely to smoke and more likely to be immigrants, which, in the United States, is associated with higher birth weight. Overall, there were similar levels of variability in risk factors for low birth weight across racial and ethnic groups in the 2 countries.

Racial and Ethnic Disparities in Low Birth Weight in the United States and England

Racial and ethnic rates of low birth weight in the United States and England are shown in Table 2. For each country, the first row shows the crude rate of low birth weight, the second row shows the percentage of low birth weight adjusted for infant gender and socioeconomic characteristics (standardized to the white mothers in each country), and the third row shows the percentage of low birth weight adjusted for gender, socioeconomic characteristics, smoking, and employment (also standardized to the white mothers).

For singleton births in the United States, the rate of low birth weight among black mothers (10.3%) was over twice that of white mothers (4.6%), whereas the rates among Hispanic (5.6%), Asian (6.4%), and American Indian (5.7%) mothers were somewhat higher than the rate for white mothers. Disparities across groups were as large in England. The rate among black mothers (9.4%) was almost twice that of white mothers (5.4%), and the rate among Asian mothers (11.5%) was the highest of all of the groups.

In the United States, adjusting for SES somewhat reduced the disparities between white mothers and those who are black, Hispanic, and American Indian. The SES-adjusted rates of low birth weight were 8.9%, 4.6%, 6.8%, and 5.0% for black, Hispanic, Asian, and American Indian mothers, respectively. Adjusting for SES did little to explain disparities in England; the SES-adjusted rates of low birth weight were 8.9% and 12.2% for black and Asian mothers, respectively. Adjustments for smoking and maternal employment also did not explain disparities in either country. SES and behavior-adjusted rates in the United States were 9.8%, 5.8%, 7.3%, and 4.5% for black, Hispanic, Asian, and American Indian mothers, respectively. The rates in England were 9.6% for black mothers and 12.9% for Asian mothers.

We conducted a corresponding set of comparisons for native-born mothers. The results are shown in Table 3. The racial and ethnic disparities in low birth weight among native-born mothers were similar to those for all mothers in each country. Again, SES explained only some of the disparities in the United States and very little of the disparities in England, and behaviors explained none of the disparities in either country. Full results for all of the mothers and native-born mothers, from logistic regressions, are shown in Appendices 1 (United States) and 2 (England).

In view of research suggesting that mean birth weights vary among ancestral and cultural groups because of differences in factors such as maternal height,29 we assessed the sensitivity of our findings to a stricter

<table>
<thead>
<tr>
<th>Variable</th>
<th>United States</th>
<th>England</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-Hispanic White</td>
<td>Non-Hispanic Black</td>
</tr>
<tr>
<td>Crude percentage of low birth weight</td>
<td>4.59</td>
<td>10.07</td>
</tr>
<tr>
<td>Percentage of low birth weight adjusted for gender and SES</td>
<td>4.59</td>
<td>8.86</td>
</tr>
<tr>
<td>Percentage of low birth weight adjusted for gender, SES, and behaviors</td>
<td>4.59</td>
<td>9.81</td>
</tr>
<tr>
<td>No.</td>
<td>3550</td>
<td>1450</td>
</tr>
</tbody>
</table>
cutoff of 2000 g. This also helped us to better distinguish “healthy” from “unhealthy” infants, because the prevalence of morbidity increases as birth weight decreases (there was insufficient sample for many of the racial and ethnic groups to use the standard 1500-g threshold for very low birth weight).3 Using the 2000-g cutoff, we obtained results that were substantively very similar to those using the conventional 2500-g threshold (Appendix 3).

We also assessed the sensitivity of our findings to a more relaxed definition of poverty in the United States, whether the mother’s household income was in the bottom 40% of the income distribution. The results were insensitive to this change (results not shown).

DISCUSSION

The findings from this study indicate that despite differences in public assistance, health care delivery systems, and immigration history, racial and ethnic disparities in low birth weight in England are as large as those in the United States and that SES, smoking, and employment explain very little of the disparities in either country.

Unlike most previous research on racial and ethnic disparities in low birth weight in the United States, the new ECLS-B data allowed us go beyond educational attainment in characterizing SES by also including measures of employment and poverty. Thus, aside from the main contribution of this article (the international comparison), we addressed an important research gap on disparities in the United States. We found that the more detailed socioeconomic data did little to explain racial and ethnic disparities in low birth weight in the United States.

The findings that racial and ethnic disparities in low birth weight are similar in magnitude in the 2 countries and that they cannot be explained by differences in SES in either country are revealing. First, they suggest that health disadvantages associated with being a minority are not a uniquely American phenomenon. They also suggest that minority health disadvantages do not solely reflect differential access to lifetime health care, although it is possible that minority status is associated with lower quality or use of care in the United Kingdom. At a minimum, our findings suggest that providing universal health care without ensuring its quality and regular use may not reduce racial and ethnic disparities in low birth weight in the United States.

Given that SES does not explain the disparities in the United States or England, the question of what does explain disparities remains. In the United Kingdom, most minorities have ancestries from countries with very high rates of low birth weight. The rates are 30% in India, 19% in Pakistan, and 30% in Bangladesh. The rates of low birth weight in commonwealth countries in the Caribbean are somewhat higher, on average, than in the United Kingdom, but there is considerable variability among them. For example, the rates in Antigua, the Bahamas, and Jamaica range from 7% to 9%, whereas the rate in Trinidad and Tobago is 23%. Rates in African countries tend to be high (eg, 18% in Senegal, 17% in Chad, 14% in Nigeria and Mozambique, and 11% in Kenya).30 Racial and ethnic disparities in low birth weight may reflect differences in biological predisposition, but it is also possible that minorities in England suffer from residual effects of poor premigration health care, which may have lasting adverse intergenerational effects on health. This potential explanation is not likely to explain the largest disparity in the United States, which is between the black and white population, because most black people have roots in the United States that go back several generations. Another potential explanation, which remains to be tested, is that minorities are disproportionately exposed to adverse environments, including social (eg, discrimination) and physical (eg, poor housing, pollution), that take a toll on their health.

Our study is subject to certain limitations. Because of the small sample sizes of minorities in other parts of the United Kingdom, we focused only on England. There was an insufficient sample to examine disparities in very low birth weight. Low birth weight is only 1 health-related outcome, and, therefore, we cannot generalize our findings to disparities in health more generally.

ACKNOWLEDGMENT

This research was supported in part by the Robert Wood Johnson Health and Society Scholars program at Columbia University.
REFERENCES


### APPENDIX 1  Multiple Logistic Regression Estimates of Associations Between Race and Ethnicity and Low Birth Weight: United States

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Mothers (N = 8250), Odds Ratio (P)</th>
<th>Native-Born Mothers (n = 5900), Odds Ratio (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>2.38 (.000)a</td>
<td>1.86 (.000)a</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1.24 (.025)b</td>
<td>1.04 (.739)</td>
</tr>
<tr>
<td>Asian</td>
<td>1.43 (.009)a</td>
<td>1.54 (.002)a</td>
</tr>
<tr>
<td>American Indian</td>
<td>1.25 (.458)</td>
<td>1.10 (.749)</td>
</tr>
<tr>
<td>Female infant</td>
<td>—</td>
<td>1.18 (.018)b</td>
</tr>
<tr>
<td>Age &lt;20 y</td>
<td>—</td>
<td>0.97 (.801)</td>
</tr>
<tr>
<td>Age ≥35 y</td>
<td>—</td>
<td>1.70 (.000)a</td>
</tr>
<tr>
<td>First birth</td>
<td>—</td>
<td>1.61 (.000)a</td>
</tr>
<tr>
<td>Married</td>
<td>—</td>
<td>0.71 (.000)a</td>
</tr>
<tr>
<td>Poor</td>
<td>—</td>
<td>1.07 (.752)</td>
</tr>
<tr>
<td>High school or GED</td>
<td>—</td>
<td>1.05 (598)</td>
</tr>
<tr>
<td>Some college</td>
<td>—</td>
<td>0.84 (.101)</td>
</tr>
<tr>
<td>College graduate</td>
<td>—</td>
<td>0.58 (.000)a</td>
</tr>
<tr>
<td>Smoked during pregnancy</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Worked during pregnancy</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Reference categories are non-Hispanic white (for race and ethnicity), 21 to 34 years (for age), and less than high school (for education). The analyses for the native-born subjects are based on the 5900 mothers in the ECLS-B analysis sample who reported in their baseline interview that they were born in the United States. — indicates that the variable was not included in the model; GED, General Education Diploma.

### APPENDIX 2  Multiple Logistic Regression Estimates of Associations Between Race and Ethnicity and Low Birth Weight: England

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Mothers (N = 10 599), Odds Ratio (P)</th>
<th>Native Born Mothers (n = 7212), Odds Ratio (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Black</td>
<td>1.83 (.000)a</td>
<td>1.69 (.001)a</td>
</tr>
<tr>
<td>Asian</td>
<td>2.29 (.000)a</td>
<td>2.64 (.000)a</td>
</tr>
<tr>
<td>Female infant</td>
<td>—</td>
<td>1.13 (.184)</td>
</tr>
<tr>
<td>Age &lt;20 y</td>
<td>—</td>
<td>0.81 (.179)</td>
</tr>
<tr>
<td>Age ≥35 y</td>
<td>—</td>
<td>1.28 (.041)b</td>
</tr>
<tr>
<td>First birth</td>
<td>—</td>
<td>1.52 (.000)a</td>
</tr>
<tr>
<td>Married</td>
<td>—</td>
<td>0.77 (.022)b</td>
</tr>
<tr>
<td>Poor</td>
<td>—</td>
<td>1.26 (.026)b</td>
</tr>
<tr>
<td>O level to less than A level</td>
<td>—</td>
<td>0.86 (.226)</td>
</tr>
<tr>
<td>A level or higher</td>
<td>—</td>
<td>0.57 (.000)a</td>
</tr>
<tr>
<td>Other qualification</td>
<td>—</td>
<td>0.47 (.002)b</td>
</tr>
<tr>
<td>Smoked during pregnancy</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Worked during pregnancy</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Reference categories are white (for race and ethnicity), 21 to 34 years (for age), and less than O level (for education). The models for native-born subjects are based on the 7212 mothers in the MCS analysis sample who completed the follow-up survey and reported at that time that they were born in the United Kingdom. — indicates that the variable was not included in the model.

### APPENDIX 3  Percentage <2000 g According to Race and Ethnicity: Singleton Births

<table>
<thead>
<tr>
<th>Variable</th>
<th>United States</th>
<th>England</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-Hispanic White</td>
<td>Non-Hispanic Black</td>
</tr>
<tr>
<td>Crude percentage of low birth weight</td>
<td>1.42</td>
<td>3.94</td>
</tr>
<tr>
<td>Percentage of low birth weight adjusted for gender and SES</td>
<td>1.42</td>
<td>3.52</td>
</tr>
<tr>
<td>Percentage of low birth weight adjusted for gender, SES, and behaviors</td>
<td>1.42</td>
<td>3.73</td>
</tr>
<tr>
<td>No.</td>
<td>3550</td>
<td>1450</td>
</tr>
</tbody>
</table>
Pertussis Booster Vaccination in HIV-Infected Children Receiving Highly Active Antiretroviral Therapy

Mark J. Abzug, MD, Lin-Ye Song, PhD, Terence Fenton, EdD, Sharon A. Nachman, MD, Myron J. Levin, MD, Howard M. Rosenblatt, MD, Stephen I. Pelton, MD, William Borkowsky, MD, Kathryn M. Edwards, MD, Jody Peters, MS, for the International Maternal Pediatric Adolescent AIDS Clinical Trials Group P1024 Protocol Team

OBJECTIVE. Our goal was to evaluate the immunogenicity and safety of pertussis booster vaccination in children infected with HIV on highly active antiretroviral therapy (HAART).

PATIENTS AND METHODS. HIV-infected children on stable HAART for ≥3 months with plasma HIV-RNA concentrations of <30 000 to 60 000 copies per mL who previously received ≥4 doses of diphtheria-tetanus-pertussis (DTP)—containing vaccine were eligible. Diphtheria-tetanus-acellular pertussis (DTaP) vaccine was administered to subjects 2 to <7 years old who had 4 previous DTP-containing vaccines, subjects 2 to <7 years old who had ≥5 previous DTP-containing vaccines and negative tetanus antibody, and subjects ≥7 to ≤13 years old who had negative tetanus antibody. Pertussis toxin and filamentous hemagglutinin antibodies were measured before and at 8, 24, and 72 weeks after DTaP vaccine.

RESULTS. Ninety-two subjects received DTaP vaccine and met criteria for analysis. Antibody concentrations were low at entry: pertussis toxin geometric mean concentration at 4.8 enzyme-linked immunosorbent assay units (EU) per mL and filamentous hemagglutinin geometric mean concentration at 4.1 EU/mL. Pertussis toxin and filamentous hemagglutinin geometric mean concentrations rose to 22.3 and 77.0 EU/mL, respectively, 8 weeks after the study DTaP vaccine. Antibody concentrations fell by 24 weeks after vaccination but remained higher than before vaccination. Predictors of response 8 weeks after DTaP vaccine included the concentration of homologous antibody, lower HIV-RNA level, and higher CD4 percentage at entry. One vaccinated subject experienced erythema and induration of ≥25 mm.

CONCLUSIONS. A DTaP vaccine booster was well tolerated by children on HAART and induced increases in antibodies. Antibody concentrations after vaccination were lower than those reported in populations uninfected by HIV. Although comparison among studies must be made with caution, these data suggest that children infected with HIV may be deficient in immunologic memory from previous DTP-containing vaccination and/or that immune reconstitution with HAART may be incomplete for pertussis antigens.
Bordetella pertussis is a cause of acute and prolonged respiratory illness in children and adults infected with HIV.1-4 The degree to which children infected with HIV are protected by immunization with pertussis-containing vaccines in early childhood has not been determined. In a small study conducted before the widespread use of highly active antiretroviral therapy (HAART), children infected with HIV had reduced antibody responses to acellular pertussis vaccine.6 Although studies suggest that HAART is associated with improved responses to revaccination with a variety of immunogens, the response of children infected with HIV who are on HAART to pertussis vaccination has not been described.7 Several reports have shown that pertussis vaccination does not have adverse effects on the clinical course, CD4 lymphocyte counts, or plasma HIV-RNA concentrations of children infected with HIV.8,9 International Maternal Pediatric Adolescent AIDS Clinical Trials Group/Pediatric AIDS Clinical Trials Group P1024 is a multicenter study designed to evaluate the immunogenicity and safety of routine pediatric vaccines in children infected with HIV who are on HAART and to assess the impact of CD4 percentage and viral load on vaccine response. This report focuses on the immunogenicity and safety of pertussis vaccination.

PATIENTS AND METHODS

Study Population

Children infected with HIV who were 2 to <19 years of age at 39 sites were eligible for enrollment if they fit into 1 of the following immunologic strata: group 1, nadir CD4 percentage before HAART of <15% and CD4 percentage at study screening of <15%; group 2, pre-HAART nadir CD4 percentage of <15% and screening CD4 percentage of ≥15%; group 3, pre-HAART nadir CD4 percentage of 15% to ≤25% and screening CD4 percentage of ≥15%; and group 4, pre-HAART nadir CD4 percentage of ≥25% and screening CD4 percentage of ≥25%. Subjects in strata 2 to 4 were required to be perinatally infected, to have been on the same HAART regimen (≥3 antiretroviral agents from ≥2 therapeutic classes) for ≥6 months, and to have a plasma HIV-RNA polymerase chain reaction (Roche Amplicor Assay; Roche Diagnostics, Indianapolis, IN) of <30 000 copies per mL. Subjects in stratum 1 were permitted to be infected via any route and were required to be receiving the same HAART regimen for ≥3 months and to have an HIV-RNA level of <60 000 copies per mL. All of the subjects were required to have previously received ≥4 doses of diphtheria-tetanus-pertussis (DTP)–containing vaccines (acellular and/or whole-cell pertussis). Exclusion criteria included receipt of other vaccines within a prescribed period preceding entry; immune suppressive or immunomodulatory therapy; blood products within 6 months; other immune diseases; and having grade 2 or higher clinical toxicities that overlap with potential vaccine-associated toxicities. Target enrollment was 300, with 75 per stratum, with a maximum of 50 subjects at ≥7 or <7 years in each stratum.

Study Protocol

Informed consent was obtained, and human experimentation guidelines of the US Department of Health and Human Services and participating institutions were followed. The age for administration of a pediatric formulation of diphtheria-tetanus-acellular pertussis (DTaP) vaccine was extended to 13 years based on satisfactory experience in other populations.10 Tetanus seronegativity at entry was used as an adjunct to select subjects eligible to receive a DTaP booster based on experience in Pediatric AIDS Clinical Trials Group 377, in which children 2 to 9 years of age infected with HIV lacking protective antibody levels to tetanus tolerated DTaP revaccination.11 DTaP (diphtheria and tetanus toxoids and acellular pertussis vaccine adsorbed: Infanrix; 25 μg of pertussis toxin [PT], 25 μg of filamentous hemagglutinin [FHA], and 8 μg of pertactin; GlaxoSmithKline Pharmaceuticals, Research Triangle Park, NC; 0.5 mL intramuscularly) was administered at the 24-week study visit to the following subjects: those at age <7 years and with 4 previous DTP vaccines; those at age ≥7 years, ≥5 previous DTP vaccines, and negative tetanus antibody; and those at age ≥7 to ≤13 years and negative tetanus antibody. Subjects received other study vaccines at other visits; responses to these will be reported separately. Therapies included among exclusion criteria for study entry were also disallowed during the study vaccination period. Subjects were monitored for adverse events for 1 hour after vaccination and by telephone 3, 7, and 14 days later. Subjects grade-3 or higher adverse reactions were evaluated within 48 hours. Specimens for pertussis antibody determination were obtained at study entry and 8 weeks after DTaP vaccination (study week 32), 24 weeks after DTaP (study week 48), and 72 weeks after DTaP (study week 96) from DTaP recipients. Subjects who did not qualify to receive DTaP based on age, previous number of DTP doses, and tetanus serology criteria defined above (“nonrecipients”) had serum for pertussis antibody measurement obtained at entry and at study weeks 48 and 96. HIV-RNA and lymphocyte subsets were determined at entry and at study weeks 24, 48, and 96.

Laboratory Assays

Tetanus immunoglobulin G was determined on duplicate serum aliquots by sandwich enzyme-linked immunosorbent assay (ELISA) (IBL-Hamburg, Hamburg, Germany). Diluted sera were incubated on plates followed by the sequential addition of enzyme-conjugated antihuman immunoglobulin G, tetramethylbenzidine substrate solution, and stop solution (0.5 mol/L of H₂SO₄).
Absorbance at 450 nm was read, and titers were determined from standard curves generated in each run using defined standards. A titer of <0.1 IU/mL was considered negative.

Antibodies to PT and FHA were determined by ELISA using a previously reported methodology. The lower limit of detection was 2 ELISA units (EU)/mL for both PT and FHA. Values of <2 EU/mL were set at 1 EU/mL for analyses.

**Statistical Analysis**

Logarithms (base 10) of PT and FHA antibody concentrations were used to normalize data where appropriate, and antibody concentrations were described with geometric mean concentrations (GMCs). Only those DTaP recipients for whom results were available at entry and 8 ± 4 weeks after DTaP administration (study week 32 ± 4 weeks) were included in analyses of immunogenicity. Their 24-week postvaccination results were included if obtained within a window of ±4 weeks (study week 48 ± 4 weeks), and their 72-week postvaccination results were included if obtained within ±12 weeks of the appropriate time point (study week 96 ± 12 weeks). Similarly, antibody concentrations were examined for DTaP nonrecipients for whom entry and study week 48 ± 4 weeks results were available; their week-96 values were included if obtained within ±12 weeks of the appropriate time point.

Univariate linear regression analyses were performed to identify predictors of entry PT and FHA antibody concentration among DTaP recipients and nonrecipients combined and for the response to pertussis booster vaccination measured 8 weeks after DTaP among vaccine recipients. Categorical variables were represented by dummy variables, and for variables having >2 categories, Tukey-Kramer simulation-based adjusted P values were used for pairwise comparisons when F tests were significant. Variables identified as at least marginally significant (P < .1) predictors of antibody concentration 8 weeks after DTaP were included in multivariate regression analyses. The strength of association between PT and FHA antibody concentrations and between entry PT and FHA antibody concentrations and change in concentration from entry to 8 weeks after DTaP were assessed with Spearman rank order methods.

**RESULTS**

**Study Population Characteristics**

A total of 263 children were enrolled; 6 were omitted from analysis because of violations of inclusion and exclusion criteria. One hundred and nineteen were eligible to receive the DTaP vaccine per protocol criteria; of these, 113 actually received DTaP. Ninety two of these had sera obtained at baseline (study entry) and 8 ± 4 weeks after DTaP; these 92 subjects comprise the primary data set of DTaP recipients. A total of 138 subjects did not qualify to receive DTaP. Nine of these were inadvertently administered DTaP. Of the remaining 129 subjects who did not receive DTaP, 98 had baseline and week 48 ± 4-week sera obtained and are included in the primary data set of DTaP nonrecipients. Age, Centers for Disease Control and Prevention clinical classification, nadir CD4 percentage before HAART, CD4 percentage at study screening, and CD4 percentage at the study DTaP visit of the primary data set of DTaP recipients differed according to immunologic strata (Table 1). Low percentages of subjects in all of the strata had received DTaP in the 2 years before study entry. Characteristics of the DTaP nonrecipients in the primary data set were generally similar to those of the DTaP recipients, with comparable relationships between immunologic strata and age, clinical classification, and CD4 measures (data not shown). Moreover, the trend toward a lower HIV-RNA concentration with ascending immune stratum was statistically significant for DTaP nonrecipients (P = .004). Because enrollment in stratum 1 was low (2 DTaP recipients and 9 nonrecipients because of a limited pool of qualifying patients), these subjects were excluded from immunogenicity analyses. There were no significant differences between subjects excluded from the primary data set and those in the primary data set with respect to baseline characteristics.

**Antibody Concentrations: Immunologic Strata Combined**

PT GMCs at entry were 4.8 EU/mL (95% confidence interval [CI]: 3.7–6.3 EU/mL) and 6.5 EU/mL (95% CI: 5.0–8.6 EU/mL) in DTaP recipients and nonrecipients, respectively (excluding stratum 1), and FHA GMCs were 4.1 EU/mL (95% CI: 2.9–5.8 EU/mL) and 7.4 EU/mL (95% CI: 5.2–10.6 EU/mL), respectively (Fig 1A and B; Table 2). Recipients of DTaP had a significant rise in PT and FHA antibody concentration 8 weeks after vaccination (study week 32), reaching GMCs of 22.3 EU/mL and 6.5 EU/mL (95% CI: 2.9–5.8 EU/mL) and 7.4 EU/mL (95% CI: 5.2–10.6 EU/mL), respectively (excluding stratum 1), and FHA GMCs were 3.7–6.3 EU/mL). Their antibody concentrations fell significantly by 24 weeks postvaccination (study week 48; decay: 55% and 57% for PT and FHA antibody, respectively), but remained higher for both FHA and PT than at entry (Table 2). A lesser but statistically significant decline in antibody concentrations occurred between weeks 24 and 72 postvaccination. The 72-week postvaccination (study week 96) GMC for FHA, but not PT, remained significantly higher than at entry. The overall decrement in antibody concentration was 70% and 65% for PT and FHA, respectively, from 8 to 72 weeks postvaccination. PT and FHA antibody concentrations among recipients of DTaP were weakly correlated at entry but were more highly correlated thereafter (Spearman correlation coefficients: 0.21 at entry [P = .05] and 0.58, 0.39, and 0.53 at 8, 24, and 72 weeks postvaccination, respectively [P ≤ .0004]).
Among nonrecipients of DTaP, PT and FHA antibody concentrations were relatively unchanged over the 96-week study period; a slight decline in PT antibody and a rise in FHA antibody from week 0 to week 48 were statistically significant, however, and the week-96 PT antibody concentration was lower than at week 0 (Fig 1 A and B; Table 2).

### Comparison of Antibody Concentrations Among Immunologic Strata

Baseline PT and FHA antibody concentrations were similar among immune strata for recipients of DTaP. Post-vaccination, strata 3 and 4 had similar PT antibody concentration profiles, with higher concentrations than those of stratum 2, although differences were not significant. A similar trend was not apparent for FHA antibody (Fig 1 C and D). Among nonrecipients of DTaP, FHA GMCs, but not PT GMCs, were directly related to immune stratum, although these findings were not significant (Fig 1 E and F).

### Antibody Concentrations and HIV-RNA Level

PT and FHA antibody concentrations at baseline were generally similar among recipients of DTaP regardless of plasma HIV viral load (Fig 2 A–D). After vaccination, PT and FHA antibody concentrations varied inversely with baseline viral load (Fig 2 A and B). Although similar patterns were evident when the HIV-RNA level at the time of DTaP administration (study week 24) was substituted for baseline viral load, especially with respect to the association of lower antibody levels with viral load of >5000 copies per mL, the relationship between viral load and antibody concentrations was not as consistent (Fig 2 C and D). This reflected some shifting of subjects among the viral load categories during the study, although total numbers of subjects in each category remained relatively stable. Among nonrecipients of DTaP, PT GMC did not vary significantly with viral load, whereas an inverse relationship between FHA GMC and baseline viral load was observed (Fig 2 E and F). The latter relationship was less consistent when the study week-24 viral load was used, although viral load of >5000 copies per mL remained associated with lower FHA GMCs (Fig 2 G and H).

### Predictors of Entry Antibody Concentration

Univariate analyses of baseline antibody concentrations among recipients of DTaP and nonrecipients combined

---

**TABLE 1** Characteristics of DTaP Recipients Included in the Primary Data Set

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All Groups</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>( p_{bc} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((N = 92)^a)</td>
<td>((n = 2)^a)</td>
<td>((n = 33)^a)</td>
<td>((n = 28)^a)</td>
<td>((n = 29)^a)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median, y</td>
<td>9.3</td>
<td>8.9</td>
<td>10.0</td>
<td>9.3</td>
<td>5.6</td>
<td>.008</td>
</tr>
<tr>
<td>( \geq 7) y, %</td>
<td>65</td>
<td>50</td>
<td>79</td>
<td>75</td>
<td>41</td>
<td>.007</td>
</tr>
<tr>
<td>Gender, male, %</td>
<td>47</td>
<td>0</td>
<td>45</td>
<td>50</td>
<td>48</td>
<td>.74</td>
</tr>
<tr>
<td>Race/ethnicity, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>14</td>
<td>10</td>
<td>.87</td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>50</td>
<td>50</td>
<td>42</td>
<td>57</td>
<td>52</td>
<td>—</td>
</tr>
<tr>
<td>Hispanic</td>
<td>38</td>
<td>50</td>
<td>45</td>
<td>29</td>
<td>38</td>
<td>—</td>
</tr>
<tr>
<td>CDC clinical classification, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N: nonsymptomatic</td>
<td>15</td>
<td>0</td>
<td>3</td>
<td>14</td>
<td>31</td>
<td>.01</td>
</tr>
<tr>
<td>A: mildly symptomatic</td>
<td>27</td>
<td>0</td>
<td>27</td>
<td>32</td>
<td>24</td>
<td>—</td>
</tr>
<tr>
<td>B: moderately symptomatic</td>
<td>36</td>
<td>50</td>
<td>30</td>
<td>39</td>
<td>38</td>
<td>—</td>
</tr>
<tr>
<td>C: severely symptomatic</td>
<td>22</td>
<td>50</td>
<td>39</td>
<td>14</td>
<td>7</td>
<td>—</td>
</tr>
<tr>
<td>Pre-HAART nadir CD4%, median</td>
<td>18</td>
<td>3</td>
<td>11</td>
<td>19</td>
<td>31</td>
<td>NA (^4)</td>
</tr>
<tr>
<td>Screening CD4%, median</td>
<td>33</td>
<td>6</td>
<td>29</td>
<td>33</td>
<td>37</td>
<td>NA (^4)</td>
</tr>
<tr>
<td>CDC immunologic classification, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III: &lt;15%</td>
<td>2</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA (^4)</td>
</tr>
<tr>
<td>II: 15% to &lt;25%</td>
<td>14</td>
<td>0</td>
<td>30</td>
<td>11</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>I: (\geq 25)%</td>
<td>84</td>
<td>0</td>
<td>70</td>
<td>89</td>
<td>100</td>
<td>—</td>
</tr>
<tr>
<td>CD4% at study DTaP visit, median</td>
<td>34</td>
<td>16</td>
<td>27</td>
<td>33</td>
<td>39</td>
<td>(&lt;.0001)</td>
</tr>
<tr>
<td>HIV-RNA concentration at study DTaP visit, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\leq 400) copies per mL</td>
<td>60</td>
<td>50</td>
<td>48</td>
<td>64</td>
<td>69</td>
<td>.62</td>
</tr>
<tr>
<td>401–5000 copies per mL</td>
<td>18</td>
<td>0</td>
<td>24</td>
<td>18</td>
<td>14</td>
<td>—</td>
</tr>
<tr>
<td>&gt;5000 copies per mL</td>
<td>22</td>
<td>50</td>
<td>27</td>
<td>18</td>
<td>17</td>
<td>—</td>
</tr>
<tr>
<td>Interval from last previous DTaP to study DTaP visit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median, y</td>
<td>5.32</td>
<td>6.60</td>
<td>5.53</td>
<td>5.59</td>
<td>3.56</td>
<td>.21</td>
</tr>
<tr>
<td>Interval (\leq 2) y, %</td>
<td>12</td>
<td>0</td>
<td>9</td>
<td>14</td>
<td>14</td>
<td>.83</td>
</tr>
</tbody>
</table>

CDC indicates centers for Disease Control and Prevention; —, no data; NA, not applicable.

\(^a\) Immunologic group 1 indicates pre-HAART nadir CD4% of \(\leq 15\%\) and screening CD4% of \(\leq 15\%\); immunologic group 2, pre-HAART nadir CD4% of \(\leq 15\%\) and screening CD4% of \(\geq 15\%\); immunologic group 3, pre-HAART nadir CD4% of \(15\%\) to \(\leq 25\%\) and screening CD4% of \(\geq 15\%\); immunologic group 4, pre-HAART CD4% of \(\geq 25\%\) and screening CD4% of \(\geq 25\%\).

\(^b\) Fisher’s exact test was used for categorical variables, and the Kruskal-Wallis test was used for continuous variables.

\(^c\) No important changes in the pattern of \( p \) values were observed when immunologic group 1 was not included.

\(^d\) Immunologic groups were defined by pre-HAART nadir CD4% and screening CD4%.

---

\( e1193 \)
FIGURE 1
A and B, PT and FHA GMCs, immunologic groups 2 to 4 combined. C and D, PT and FHA GMCs of DTaP recipients divided among immunologic groups; DTaP given at study week 24 and antibody measured at entry and 8, 24, 72 weeks postvaccination (study weeks 0, 32, 48, and 96). E and F, PT and FHA GMCs of DTaP nonrecipients divided among immunologic groups; antibody measured at study weeks 0, 48, and 96. Error bars indicate 95% CIs. GMCs were compared with F tests (general linear models procedure); immunologic group 1 is presented in C through F but was not included in comparisons.
(excluding an outlier with an extreme elevation of PT but not FHA antibody at entry) demonstrated an inverse relationship between FHA antibody and the interval between the last previous pertussis vaccination and study entry ($P = .04$), a higher PT GMC in boys ($P = .03$), a higher PT GMC in subjects 2 to <7 years of age ($P = .05$), a higher FHA GMC in the subjects who did not qualify to receive study DTaP ($P = .02$), and a direct relationship between baseline PT and FHA antibody concentrations (Spearman correlation coefficient: 0.40; $P < .0001$). A direct relationship between baseline FHA antibody concentration and entry CD4 percentage was also observed (Spearman correlation coefficient: 0.20; $P = .006$). The entry FHA GMC of subjects whose viral load closest to the most recent previous pertussis vaccination was >5000 copies per mL was lower than the GMC of subjects with viral loads of >400 to ≤5000 copies per mL (3.5 vs 21.1 EU/mL; $P = .002$) and marginally lower than that of subjects with viral loads of ≤400 copies per mL (3.5 vs 8.8 EU/mL; $P = .06$), although these viral load data were lacking for ≥50% of subjects. Neither baseline PT nor FHA antibody concentrations correlated with race, nadir CD4 percentage before HAART, lowest CD4 percentage ever, CD4 percentage at initiation of HAART, duration of the entry HAART regimen, CD4 percentage nearest to the most recent previous pertussis vaccination, or entry lymphocyte count, CD19%, immune stratum, or viral load.

**Predictors of Response to Study Vaccination**

Univariate analyses demonstrated that PT and FHA antibody concentrations 8 weeks after DTaP booster (study week 32) were directly related to the concentration of the homologous antibody at entry (Tables 3 and 4). The change in concentration of FHA but not PT antibody from baseline to 8 weeks postvaccination was also positively related to the homologous antibody concentration at entry (Spearman correlation coefficient: 0.45; $P < .0001$). Eight weeks after vaccination, GMCs were not significantly related to immunologic stratum, lowest CD4 percentage before initiation of the first HAART regimen, or lowest CD4 percentage ever. However, PT antibody was directly related to CD4 percentage at entry and at the time of study DTaP administration. Both PT and FHA GMCs 8 weeks after DTaP were inversely related to baseline viral load (Table 3); pairwise comparisons demonstrated significant differences in PT and FHA antibodies between subjects with viral loads of ≤400 copies per mL versus subjects with viral loads of >5000 copies per mL ($P = .05$). Although similar trends were observed in relation to the viral load at the time of DTaP administration (study week 24), they did not approach significance. Scatter plots of the logarithm of the 8-week postvaccination PT and FHA antibody concentrations versus the logarithm of viral load at baseline or at the DTaP visit for subjects with viral loads of >400 copies per mL suggested negative slopes but with considerable scatter and no HIV-RNA thresholds predictive of greatly diminished responses (data not shown).

In multivariate analyses with antibody concentration 8 weeks after vaccination as the outcome and predictors consisting of the concentration of homologous antibody at entry and baseline viral load, the former predictor remained highly significant ($P = .0001$), whereas entry HIV-RNA concentration was only marginally related ($P = .09$) for both PT and FHA antibodies. In a second multivariate analysis, PT antibody 8 weeks after vaccination was significantly associated with baseline PT antibody concentration ($P < .0001$) and CD4 percentage at the time of study DTaP administration ($P = .03$) but not with race ($P = .2$) or baseline viral load ($P = .3$).

**Safety**

One subject had a grade-3 event, localized erythema and induration (≥25 mm but <50% of involved extremity), judged related to DTaP. This represented 1 (0.8%) of 122 subjects who received study DTaP; 1 (2.4%) of 41 subjects in stratum 2 who received study DTaP; and 1 (10%) of 10 subjects <7 years of age who had received ≥5 doses of DTaP before entry, had negative tetanus titers, and received study DTaP. None of the 9 subjects who did not qualify to receive DTaP but were inadvertently ad-

---

**TABLE 2**

GMCs of PT and FHA Antibody in Recipients of DTaP and Nonrecipients at Each Study Visit

<table>
<thead>
<tr>
<th>Study Week (Weeks After Booster Vaccine for DTaP Recipients)$^a$</th>
<th>DTaP Recipients$^b$</th>
<th>DTaP Nonrecipients$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N$^c$</td>
<td>PT, GMC (95% CI), EU/mL</td>
</tr>
<tr>
<td>0</td>
<td>90</td>
<td>4.8 (3.7–6.3)</td>
</tr>
<tr>
<td>32 (8)</td>
<td>90</td>
<td>22.3 (15.9–31.2)$^{de}$</td>
</tr>
<tr>
<td>48 (24)</td>
<td>81</td>
<td>10.1 (7.1–14.4)$^{de}$</td>
</tr>
<tr>
<td>96 (72)</td>
<td>78</td>
<td>6.8 (5.1–9.1)$^d$</td>
</tr>
</tbody>
</table>

$^a$ DTaP recipients were administered a booster vaccine at study week 24.

$^b$ Immunologic groups 2, 3, and 4 are combined; subjects in immunologic group 1 (2 DTaP recipients and 9 DTaP nonrecipients) are excluded.

$^c$ Week-32, -48, and -96 serologies for subjects in the primary data set were included if they were obtained within windows of ±4, ±4, and ±12 weeks, respectively.

$^d$ P ≤ .03 in a comparison of GMC with GMC at the previous time point by matched-pairs Wilcoxon signed-rank test.

$^e$ P ≤ .05 in a comparison of GMC with GMC at week 0 by matched pairs Wilcoxon signed-rank test.

$^f$ NA indicates not applicable; week-32 serology was not obtained per the protocol for DTaP nonrecipients.
FIGURE 2
A–D, PT and FHA GMCs of DTaP recipients according to HIV-RNA concentration at baseline and at DTaP visit (study week 24); antibody measured at entry and 8, 24, 72 weeks postvaccination (study weeks 0, 32, 48, and 96). E–H, PT and FHA GMCs of DTaP nonrecipients according to viral load at baseline and at week 24; antibody measured at study weeks 0, 48, 96. Error bars indicate 95% CIs. GMCs were compared with F tests (general linear models procedure); immunologic group 1 was not included in graphs or comparisons.
ministered vaccine suffered a grade-3 reaction. There were no grade-4 events related to study vaccinations, and no life-threatening adverse events or deaths occurred. There was no significant change during the study in the percentage of recipients of DTaP with plasma HIV-RNA levels of ≤400 copies per mL (60%–61% at each visit) or in the mean CD4 percentage after DTaP administration (33%–34% at each visit).

DISCUSSION

Interpretation of serologic responses to pertussis vaccination is challenging, because correlates of protection have not been established, there probably are contributions of multiple antibodies and cellular immunity to protection, and standardized functional assays are not available. Although antibody concentrations to antigens such as PT or to a combination of antigens correlated with clinical protection in some studies, they did not in others.13–17 Nevertheless, several conclusions can be drawn by comparing the antibody levels that we observed with levels in HIV-uninfected populations. Baseline antibody concentrations in our cohort were low, similar to or lower than concentrations in 15- to 20-month-olds who had received a 3-dose primary series of acellular pertussis vaccines in infancy,18 4- to 6-year-olds who had received a primary series in infancy and a booster dose at 12 to 18 months,16,17 and adolescents and

### TABLE 3

<table>
<thead>
<tr>
<th>Categorical Variable</th>
<th>N</th>
<th>PT</th>
<th>FHA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immunologic group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nadir CD4% &lt; 15% and screening CD4% ≥ 15%</td>
<td>33</td>
<td>.27</td>
<td>15.53</td>
</tr>
<tr>
<td>Nadir CD4% 15% to ≤25% and screening CD4% ≥ 15%</td>
<td>28</td>
<td>27.17</td>
<td>.74</td>
</tr>
<tr>
<td>Nadir CD4% ≥ 25% and screening CD4% ≥ 25%</td>
<td>29</td>
<td>27.79</td>
<td>97.43</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;7 y</td>
<td>31</td>
<td>.19</td>
<td>30.21</td>
</tr>
<tr>
<td>≥7 y</td>
<td>59</td>
<td>19.00</td>
<td>74.73</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>43</td>
<td>.72</td>
<td>23.76</td>
</tr>
<tr>
<td>Female</td>
<td>47</td>
<td>21.03</td>
<td>76.34</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>11</td>
<td>.08</td>
<td>39.10</td>
</tr>
<tr>
<td>Black</td>
<td>45</td>
<td>27.48</td>
<td>85.87</td>
</tr>
<tr>
<td>Hispanic</td>
<td>34</td>
<td>14.10</td>
<td>50.05</td>
</tr>
<tr>
<td>RNA group at entry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤400 copies per mL</td>
<td>53</td>
<td>.06</td>
<td>28.65</td>
</tr>
<tr>
<td>401–5000 copies per mL</td>
<td>18</td>
<td>23.64</td>
<td>63.44</td>
</tr>
<tr>
<td>&gt;5000 copies per mL</td>
<td>19</td>
<td>10.48</td>
<td>32.32</td>
</tr>
<tr>
<td>RNA group at study DTaP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤400 copies per mL</td>
<td>54</td>
<td>.35</td>
<td>26.25</td>
</tr>
<tr>
<td>401–5000 copies per mL</td>
<td>17</td>
<td>22.14</td>
<td>112.07</td>
</tr>
<tr>
<td>&gt;5000 copies per mL</td>
<td>19</td>
<td>14.10</td>
<td>41.57</td>
</tr>
</tbody>
</table>

* Data are from an F test from a general linear models procedure.

### TABLE 4

<table>
<thead>
<tr>
<th>Continuous Variable</th>
<th>N</th>
<th>PT</th>
<th>FHA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of entry HAART regimen</td>
<td>90</td>
<td>.26</td>
<td>-0.0010b</td>
</tr>
<tr>
<td>Homologous antibody concentration (log10) at entry</td>
<td>90</td>
<td>.52</td>
<td>0.0011b</td>
</tr>
<tr>
<td>Nadir CD4% before HAART</td>
<td>90</td>
<td>&lt;0.0001</td>
<td>0.54b</td>
</tr>
<tr>
<td>CD4% at entry</td>
<td>90</td>
<td>0.17</td>
<td>0.0096b</td>
</tr>
<tr>
<td>Total lymphocyte count at study DTaP</td>
<td>79c</td>
<td>0.05</td>
<td>0.017b</td>
</tr>
<tr>
<td>CD4% at study DTaP</td>
<td>90</td>
<td>.27</td>
<td>-0.00010b</td>
</tr>
<tr>
<td>CD19% at study DTaP</td>
<td>90</td>
<td>0.03</td>
<td>0.019b</td>
</tr>
</tbody>
</table>

* Data are from an F test from a general linear models procedure.
* Data are regression parameters.
* Data were not available for some subjects.
adults tested before receiving a booster dose\textsuperscript{19–21} (PT GMCs = 2–14 EU/mL and FHA GMCs = 0–62 EU/mL after various vaccine formulations). The low baseline antibody concentrations in our population, despite \( \geq 4 \) previous pertussis vaccinations, were also consistent with low levels found in a cross-sectional analysis of 10- to 49-year-olds in the Third National Health and Nutrition Examination Survey. In this large serosurvey of a nationally representative sample of adolescents and adults remote from vaccination whose specimens were tested in the same laboratory used in the present study, PT antibody concentrations were \(<1\) EU/mL in 16\% and \(<20\) EU/mL in 85\%, and FHA antibody concentrations were \(<1\) EU/mL in 1\% and \(<20\) EU/mL in 58\%.\textsuperscript{12}

After booster vaccination, our subjects had significant increases in PT and FHA antibody concentrations, although levels achieved were generally less than in children uninfected with HIV after a primary series of acellular pertussis vaccines in infancy,\textsuperscript{22–24} after a first booster dose of DTaP at 12 to 28 months of age preceded by primary vaccination in infancy,\textsuperscript{16,25} and after a second booster dose of DTaP at 4 to 6 years\textsuperscript{16} (PT GMCs: 51–137 EU/mL and FHA GMCs: 103–650 EU/mL after Infanrix or different vaccine formulations containing the same concentrations of PT and FHA antigens). The antibody concentrations our cohort attained were also lower than postvaccination antibody concentrations reported for acellular pertussis vaccines containing approximately one third the amount of PT and FHA antigens (ie, 8 \( \mu \)g of PT and 8 \( \mu \)g of FHA) in HIV-uninfected 10- to 18-year-olds\textsuperscript{20} and adults\textsuperscript{21,26,27} (PT GMCs: 38–140 EU/mL and FHA GMCs: 354–906 EU/mL). Direct comparison among studies must be done with caution because of differences in populations, vaccines, study design, and serologic assays.\textsuperscript{17} For example, a limitation of the present study is that serologic response was assessed 2 months after vaccination (range: 1–3 months) rather than 1 month postvaccination as was done in most comparator studies. In addition, we did not measure the antibody to pertactin because this assay was not included in the assay standardization reported for the laboratory used in the present study.\textsuperscript{12} Despite these limitations, previously described populations uninfected with HIV seemed to mount higher antibody responses after boosting than did our subjects infected with HIV. Our use of tetanus seronegativity as a screen for administering DTaP to subjects \(<7\) years of age who had previously received \( \geq 5 \) DTP vaccines and for subjects 7 to 13 years of age may have biased the vaccinated group in favor of a less immune-competent subset. However, the similar profiles of subjects who did and did not qualify to receive DTaP argue against significant bias. On the other hand, this selection strategy may have contributed to the low rate of adverse events that we observed with pediatric DTaP in children infected with HIV up to 13 years of age. This low reactogenicity was similar to that observed with primary vaccination in infants\textsuperscript{28} and contrasts with rates of moderate-severe erythema, swelling, and pain of \( \geq 26\% \) in 4- to 6-year-old children uninfected with HIV who were receiving a fifth dose of Infanrix.\textsuperscript{29}

The rapid rates of antibody decay were similar to rates of decline after pertussis vaccination in infants and young children (\( >50\% \) in 6 months\textsuperscript{12,15}; 85\%–98\% over 4 years\textsuperscript{16}) and adolescents and adults (42\%–58\% between 1 and 6 months postvaccination; 61\%–73\% between 1 and 18 months postvaccination; and 80\%–86\% over 3 years postvaccination) uninfected with HIV.\textsuperscript{19} Reported rates of decay of PT antibody in the year after vaccination of 52\% to 78\% mirror the decrement in cell-mediated immune responses after vaccination and the fall in titers in the year after natural \textit{B pertussis} infection (66\% for PT and 19\% for FHA).\textsuperscript{19,21,30} Despite waning antibody titers, subjects uninfected with HIV rapidly develop high PT and FHA antibody levels after booster vaccination, consistent with an anamnestic response and immunologic memory.\textsuperscript{16} Whether this is true of HIV-infected populations is unknown.

De Martino et al\textsuperscript{6} demonstrated that 9 of 12 children infected with HIV receiving monotherapy (zidovudine or didanosine) who were given 3 doses of acellular pertussis vaccine had a fourfold or more rise in antibody to \( \geq 1 \) of PT, FHA, or pertactin. Antibody titers 2 months after the third dose were significantly lower than those of control subjects uninfected with HIV: PT GMC was 30 vs 106 EU/mL, and FHA GMC was 26 vs 760 EU/mL. Although our study did not directly compare immune responses in children infected with HIV on HAART with children uninfected by HIV, postbooster GMCs attained by our population were more similar to the postvaccination titers of subjects infected with HIV who were on monotherapy than the HIV-uninfected control subjects in the study by De Martino et al.\textsuperscript{6} This may reflect a deficiency of immunologic memory induced by previous DTP vaccination and absence of an anamnestic response to boosting. Alternatively, the low pertussis antibody response, which contrasts with ELISA antibody responses to a pneumococcal conjugate plus polysaccharide vaccine series that were quite similar to those of children uninfected with HIV,\textsuperscript{31} may be indicative of immunogen-specific variability in the degree of immune reconstitution that accompanies HAART. Deficient pathogen-specific immune responses have been documented in patients infected with HIV who were on HAART.\textsuperscript{32–35} Similarly, although small studies of children treated with HAART demonstrated improved antibody responses to measles, mumps, rubella, tetanus, and \textit{Haemophilus influenzae} type b vaccines, other reports found that children and adults infected with HIV and on HAART have limited antibody and cellular responses to hepatitis A and tetanus vaccines.\textsuperscript{7,11,36–39} Whether additional doses of DTaP or higher doses of pertussis antigens while on HAART would stimulate antibody levels similar to
those seen in children uninfected with HIV remains to be tested.39

Antibody concentration at entry was a strong predictor of response to booster vaccine, and entry FHA antibody concentration correlated with change in FHA antibody concentration from before to after booster vaccination, despite the bias in the negative direction of regression toward the mean. This is similar to observations in an adolescent and adult acellular pertussis vaccine study21 and in our study of pneumococcal conjugate and polysaccharide vaccines in children infected with HIV.31 It suggests that baseline antibody concentration may be a marker of immune competence, perhaps reflecting genetic factors that influence vaccine responsiveness.40-43 Baseline PT antibody concentration did not correlate significantly with change in PT antibody concentration after booster vaccine, possibly because the magnitude of PT antibody responses was less robust overall.

PT antibody response correlated with the CD4 percentage at study entry and at the time that the booster DTaP dose was administered but not with the nadir CD4 percentage before HAART. This is consistent with other studies in HIV-infected populations that showed direct relationships between the CD4 percentage and response to a variety of vaccines, including pertussis-containing vaccine.8,37,44-50 It suggests that changes in the CD4 percentage in response to HAART are more predictive of vaccine response than the CD4 percentage before initiation of HAART, as we observed with pneumococcal vaccination.31 PT and FHA antibody responses were inversely related to HIV-RNA level, also similar to our findings with pneumococcal vaccination11 and to observations with measles and varicella vaccines.38,51 Whether ongoing HIV replication in the face of HAART reflects antiretroviral resistance, medication nonadherence, or pharmacologic factors, these findings are consistent with observed HIV-induced impairment of B-cell–CD4-cell interactions.52

CONCLUSIONS
The relatively low pertussis antibody responses to boosting in children infected with HIV suggest that a single booster dose of pertussis vaccine, even in the context of HAART, may be insufficient to induce an immune response comparable to that of children, adolescents, and adults who are uninfected with HIV. Although minimum protective levels of antibody to pertussis antigens have not been defined, these findings further suggest that immunologic memory from previous pertussis vaccination may be lacking in children infected with HIV and/or that immune reconstitution with HAART may be incomplete with respect to pertussis antigens. Further investigation is needed to define the optimal approach to preventing pertussis infection in children infected with HIV and to better understand the immunologic deficits that may be present in children infected with HIV despite treatment with HAART.

ACKNOWLEDGMENTS
This work was supported in part by the International Maternal Pediatric Adolescent AIDS Clinical Trials Group of the National Institute of Allergy and Infectious Disease (NIAID) and the Pediatric/Perinatal HIV Clinical Trials Network of the National Institute of Child Health and Human Development (NICHD) (NIAID sites were supported by NIAID grant U01 AI068632, and NICHD sites were supported by NICHD contract N01-HD-3-3345). This work was also supported in part by the General Clinical Research Center Units, funded by the National Center for Research Resources, National Institutes of Health. GlaxoSmithKline Pharmaceuticals provided study vaccine.

Additional members of the P1024 Protocol Team include Shirley Jankelevich, MD (Pediatric Medicine Branch, National Institute of Allergy and Infectious Diseases, Bethesda, MD), Jennifer Read, MD (Pediatric, Adolescent, and Maternal AIDS Branch, National Institute of Child Health and Human Development, Bethesda, MD), Victoria Hadhazy, MA (International Maternal Pediatric Adolescent AIDS Clinical Trials Group/Pediatric AIDS Clinical Trials Group Operations Office (Rockville, MD), Alison Robbins, MA (International Maternal Pediatric Adolescent AIDS Clinical Trials Group/Pediatric AIDS Clinical Trials Group Operations Office, Rockville, MD), Jane C. Lindsey, ScD (Statistical and Data Analysis Center, Harvard School of Public Health, Boston, MA), Janice Hodge, RN, BS (Frontier Science and Technology Research Foundation, Amherst, NY), Dorothy Smith, MS, CPNP (University of Massachusetts Memorial Medical Center, Worcester, MA), Paul Tran, RPh (Pharmaceutical Affairs Branch, National Institute of Allergy and Infectious Diseases, Bethesda, MD), Grace Aldrovandi, MD (Children’s Hospital of Los Angeles, Los Angeles, CA), Rachel Barrett, BS (University of Colorado School of Medicine, Denver, CO), William Kabat, BS (Chicago Children’s Memorial Hospital, Chicago, IL), Velma Keeley, BS (Wyeth-Ayerst Global Pharmaceuticals, St Davis, PA), Paul Willems, MD (GlaxoSmithKline Pharmaceuticals, Collegeville, PA), and Carol Gore (Panorama Village, TX).

Participating sites and site personnel include Chicago Children’s Memorial Hospital (Ram Yoge, MD), University of Medicine and Dentistry of New Jersey-New Jersey Medical School (Barry Dashefsky, MD, Linda Bettica, RN, and Paul Palumbo, MD), Harlem Hospital (Elaine Abrams, MD, Maxine Frere, RN, and Lisa Gaye Robinson, MD), Metropolitan Hospital Center (Mahr-ukh Bamji, MD), Long Beach Memorial Hospital (Audra Deveikis, Susan Marks, Karen Elkins, and Lisa Melton), San Juan City Hospital (Eleanor Jimenez, MD), Los Angeles County Medical Center (James Homans, MD, Ana
REFERENCES


7. Melvin AJ, Mohan KM. Response to immunization with measles, tetanus, and Haemophilus influenzae type b vaccines in children who have human immunodeficiency virus type 1 infection and are treated with highly active antiretroviral therapy. Pediatrics. 2003;111(6). Available at: www.pediatrics.org/cgi/content/full/111/6/e641


Adherence to Nasal Positive Airway Pressure Therapy Among School-aged Children and Adolescents With Obstructive Sleep Apnea Syndrome

Elizabeth C. Uong, MD*, Mary Epperson, RN*, Sharon A. Bathon, RN*, Donna B. Jeffe, PhD*

*Division of Allergy and Pulmonary Medicine, Department of Pediatrics, and Division of Health Behavior Research, Department of Medicine, Washington University School of Medicine, St Louis, Missouri

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVES. Although many children with obstructive sleep apnea syndrome have complete resolution of obstructive sleep apnea syndrome after adenotonsillectomy, some patients have persistent obstructive sleep apnea syndrome requiring positive airway pressure treatment. Little is known about positive airway pressure adherence among school-aged children and adolescents.

PATIENTS AND METHODS. We retrospectively reviewed records from January 2000 through December 2004 to assess positive airway pressure adherence following a comprehensive patient- and parent-focused positive airway pressure education program for children 7 to 19 years of age with persistent obstructive sleep apnea syndrome subsequent to indicated adenotonsillectomy. A polysomnogram was obtained before and after initiation of positive airway pressure therapy. Adherence was defined as >4 hours per night and ≥5 nights per week of positive airway pressure use. Clock-counter meters determined hours per night and nights per week of positive airway pressure use; parents estimated hours per night of positive airway pressure use. Nonparametric tests assessed associations between adherence and various clinical parameters and symptoms.

RESULTS. Forty-six patients (56% male; 39% black, 61% white; mean age: 13.6 years; mean BMI: 39.8 kg/m²) were included. Two refused positive airway pressure. Meter readings were available for 27 patients (59%); positive airway pressure was used, on average, 7.0 hours per night, 73% of the week, and for a mean of 18.1 months. Nineteen (70%) were adherent regardless of age. There was good agreement between parental report and meter readings. Patients with greater improvement in apnea-hypopnea index were more likely to be adherent. Clinical parameters and symptoms improved after positive airway pressure therapy regardless of age or adherence.

CONCLUSIONS. In this retrospective study, positive airway pressure adherence and symptom improvement among school-aged children and adolescents was achieved.
with comprehensive patient and parent education and follow-up.

Untreated obstructive sleep apnea syndrome (OSAS) in children can result in serious health complications. It imposes a substantial health care burden, along with a negative effect on health-related quality of life. Although adenotonsillectomy remains the treatment of choice for most cases of OSAS in children, a cumulative cure rate of only 80% is achieved, with a percentage of patients with persistent OSAS requiring positive airway pressure (PAP), using single-level or bi-level airway pressure.

PAP therapy is effective for the long-term treatment of OSAS both in adults and children. Nonadherence to PAP use, however, limits its effectiveness. In adults, PAP adherence has been reported to vary between 65% and 80%, with 8% to 15% of patients refusing this treatment after a single night’s use in the laboratory. The definition of adherence varies across studies, however. Although many studies have focused on nightly duration of PAP use, night-to-night consistency and frequency of use may be as important. One study, for example, reported that only 46% of adult patients used PAP for ≥4 hours on 70% of the days.

Studies in children are even more varied, with adherence estimated to be between 50% and 100%. These estimates, however, have been based on parental reports or questionnaires sent to pediatric practitioners. Two recent studies measured objective PAP adherence in children using clock-counter/meter readings. Nevertheless, all of these studies included patients from a wide age range of 6 months to 18 years, with adolescents reported to be the least adherent. One might argue that adherence in younger children is primarily a function of parent cooperation. Adherence by school-aged children and adolescents, however, may need to be examined against a background of changing physical, psychological, and social development unique to this group.

Studies in adults suggest that high PAP adherence rates can be achieved in an environment that fosters patient education, comprehensive follow-up, and integrated care. Variables found to be significantly correlated with PAP usage in adults include higher apnea-hypopnea index (AHI), higher Epworth sleepiness scores, higher baseline BMI, male gender, older age, and patient support and education. Only 1 similar study has been conducted in children, and no similar studies have been conducted on school-aged children and adolescents, the group commonly considered to be least adherent. In a prospective study, we describe nasal PAP effectiveness and adherence among school-aged children and adolescents who had been followed in our clinic through a comprehensive program dedicated to PAP education and follow-up.

Methods

Participants

The study was approved by the institutional review board at Washington University School of Medicine. Patient records from the St Louis Children’s Hospital Sleep Clinic beginning January 2000 through December 2004 were retrospectively reviewed. We included all of the children ≥7 years of age who, despite indicated adenotonsillectomy, had OSAS confirmed by polysomnogram based on published standard criteria. The patient was considered to have OSAS if an AHI ≥5 and/or presence of any abnormalities in gas exchange was documented, including a pulse oxygen saturation ($\text{SpO}_2$) nadir of ≤85%, $\text{SpO}_2$ <90% for ≥10% of total sleep time (TST), peak end-tidal CO$_2$ (ETCO$_2$) ≥55 mm Hg, and ETCO$_2$ >50 mm Hg for ≥10% of TST. All of the patients underwent a baseline polysomnogram evaluation followed by a polysomnogram PAP titration study. Patients with severe cognitive impairment (who, therefore, were unable to understand or adhere to instructions), who were on PAP therapy for indications other than OSAS, or who had received PAP treatment before January 2000 were excluded.

Procedures

Patients were assessed in a dedicated sleep clinic by 1 physician, who is a pediatric, board-certified sleep specialist. A set of standard questions was used to determine baseline clinical symptoms observed by parents; this included the presence of snoring, witnessed episodes of apnea and choking respirations during sleep, sleeping patterns, presence of secondary enuresis, and subjective assessment of changes in school performance, behavior, and sleepiness since the onset of nighttime symptoms. Patients and their parents were also given a 30-minute educational session about OSAS as part of their initial clinic visit. Health consequences of OSAS and treatment options, including an overview of PAP therapy, were reviewed before the patients’ baseline polysomnogram.

All of the patients underwent baseline polysomnogram evaluation. After baseline data were collected, a brief PAP trial was offered during the last hour of the initial overnight polysomnogram to familiarize patients with PAP therapy. Baseline polysomnogram evaluation results were then discussed with the patient and his or her parents 2 weeks after the initial polysomnogram, and additional PAP education with emphasis on the importance of adherence was provided. A second polysomnogram was obtained to determine optimal PAP pressure settings. Follow-up telephone calls by a dedicated nurse were made shortly after optimal PAP settings were in place, and patients had clinic follow-up visits 2
to 4 weeks later and every 6 months thereafter. Problem areas, if any, were determined at each follow-up visit. The number of hours of PAP use was obtained from the patient and/or parent and, for those patients who used a clock-counter/meter, the number of hours and frequency of use during the week were obtained from the meter readings. Patients with meter readings had PAP devices with a built-in monitoring chip that registered use when the set pressure was maintained and then collected and stored the PAP usage data. Monitored data were downloaded using the Encore Pro Data Management System (Respironics, Murrysville, PA) during the clinic follow-up visit. Detailed meter-reading summaries of the hours of daily use, time of days used, and days used per month were reviewed.

Polysomnography and Institution of PAP Therapy

All of the polysomnograms were performed at the St Louis Children’s Hospital sleep laboratory. Patients fell asleep without the aid of sleep medication and were observed for \( \geq 8 \) hours overnight in the company of \( \geq 1 \) parent. Parameters measured included chest and abdominal wall movement by piezoelectric respiratory belts (Respironics); heart rate by electrocardiogram; and airflow with sidestream end-tidal capnograph, which provided breath-by-breath assessment of ETCO\(_2\) levels (Capnocheck Plus; BCI International, Waukesha, WI) and an oronasal thermistor (Respironics). Arterial oxygen saturation was assessed using pulse oximetry (SpO\(_2\)) with simultaneous recording of the pulse waveform (BCI International). Bilateral electrooculogram, 8 channels of electroencephalogram, chin electromyogram, and analog output from a body-position sensor were monitored. Tracheal sound was monitored with a microphone sensor (Respironics), and synchronized video recording was performed. All of the measures were digitized using a commercially available polysomnographic system (Alice III; Respironics), where data were acquired, recorded, and stored. Raw data were manually scored by 30-second epoch, according to published standards.\(^{20}\) Obstructive apnea was defined as the absence of airflow with continued chest and abdominal wall movement for a duration of 2 breath cycles.\(^{20}\) Hypopnea was defined as a \( \geq 50\% \) decrease in nasal airflow with a corresponding hypoxemia and/or arousal.\(^{20}\) The AHI was defined as the number of apneas and hypopneas per hour of TST. The mean SpO\(_2\) and SpO\(_2\) nadir were determined. Arousals were defined as recommended by the American Sleep Disorders Task Force Report.\(^{31}\)

A titration polysomnogram was obtained using the same polysomnogram parameters as described. The goal of the titration polysomnogram was to improve gas exchange and to reduce the AHI to normal or near normal levels through PAP at a level tolerated by the patient. PAP was started at single level pressure of 5 cm H\(_2\)O and increased by increments of 2 cm H\(_2\)O when needed. Patients who seemed uncomfortable or who required single level pressures of \( \geq 15 \) cm H\(_2\)O were switched to bilevel pressures, beginning at 10/5 with \( \geq 5 \) cm H\(_2\)O difference between inspiratory and expiratory pressures. Once optimal pressures were determined, the family received a follow-up telephone call from a dedicated sleep nurse to review study results, PAP pressures, and instructions regarding home health PAP set-up and clinic follow-up. Humidifiers were used for all of the patients. A representative of the home health care company who visited the patients in their home offered various masks for the best fit, and follow-up appointments in the sleep clinic were arranged 2 to 4 weeks into PAP therapy.

Measures

Assessment of PAP Effectiveness

Objective polysomnogram findings, including AHI, arterial SpO\(_2\) nadir, mean ETCO\(_2\) levels, and presence of obstructive hypoventilation (defined as the presence of ETCO\(_2\) of \( \geq 50 \) mm Hg for \( >10\% \) of TST), were compared before PAP and after optimal PAP settings were initiated on repeat polysomnogram. The repeat polysomnogram was performed on PAP. In addition, clinical symptoms based on history, including nocturnal (snoring, witnessed apneas, choking episodes, and enuresis) and daytime symptoms (observed daytime sleepiness, hyperactivity, and school performance), were compared before and after home PAP use. A symptom score was computed on the basis of the total number of 4 nocturnal symptoms and daytime somnolence reported, for a total possible symptom score of 5.

Assessment of PAP Adherence

Both objective (clock-counter/meter) and subjective (parental report) data were obtained only from the patients with meter readings to determine the mean number of hours of PAP use per night and the frequency of PAP use during the week (number of nights per week in which PAP was used). Patients were considered to be adherent if the meter readings indicated PAP use of \( >4 \) hours per night of use and \( \geq 5 \) nights per week, similar to the calculation of adherence in some adult studies.\(^{19,32}\)

Statistical Analysis

Nonparametric tests (eg, \( \chi^2 \) tests and Mann-Whitney \( U \) tests) were used to measure the significance of univariate associations between adherence to PAP use (for the 27 patients with meter readings) and each of the clinical outcomes, presence and absence of individual symptoms, total number of symptoms, and change in symptoms, as well as gender, race, age, and BMI. Spearman correlations measured associations among continuous variables, such as age, average hours of PAP use per night, total number of symptoms before and after PAP,
and change in symptoms. We illustrated the agreement between parent-reported and meter readings of mean hours per night of PAP use (for patients who used clock-counter meters) using a Bland-Altman plot, which plots the differences between the 2 measures against the averages of the 2 measures. Wilcoxon signed-rank tests were used to measure the significance of changes in each of the symptoms and in the total number of symptoms measured before and after home PAP use. We also ran a repeated-measures analysis of variance to test whether the change in the total number of symptoms was different between adherent and nonadherent patients.

RESULTS

Patient Characteristics

Forty-six patients who were seen in the St Louis Children’s Hospital Sleep Clinic from January 2000 through December 2004 met the inclusion criteria (Table 1). No patients were excluded from the study on the basis of the cognitive-impairment exclusion criterion. The mean age of these patients was 13.6 years (SD: 3.1 years), and mean BMI was 39.8 kg/m² (SD: 15.2 kg/m²), ranging from 16.7 to 76.4 kg/m² (BMI z scores ranged from −1.52 to 2.41). As shown in Table 1, more than half of the sample was <15 years of age. Thirty-one patients (67%) had obesity as an associated comorbidity of OSAS. Mean AHI at baseline was 28.4 events per hour. Of the 46 patients, 2 patients (4%), aged 11 and 19 years, completely refused PAP therapy despite further education and follow-up. These 2 patients had very mild OSA on the basis of a baseline AHI of 1.5 and 6.4 events per hour, respectively. Of the 44 patients on PAP, 30 (68%) were on single pressure with a mean continuous PAP pressure of 8.8 cm H₂O (SD: 2.1 cm H₂O); and 14 (32%) were on bilevel pressures, with a mean inspiratory pressure of 14.9 (SD: 3.5) and a mean expiratory pressure of 8.4 (SD: 2.6). Eleven patients reported problems during PAP use. The most common problem involved mask fit (8 of 11); other problems included skin breakdown, dry nose, and stuffy ears, each reported by 1 patient.

Twenty-seven patients (59%) had clock-counter/meter readings. At the time the data were collected by chart review, the average length of home PAP use among the 27 patients using PAP with meters had been 18.1 months (SD: 11.8 months; range: 3–43 months; median: 15.0 months).

PAP Effectiveness

As shown in Tables 2 and 3, changes in AHI, SpO₂ nadir, and clinical symptoms showed significant improvement regardless of PAP mode (single or bilevel pressure). Polysomnogram parameters showed significant improve-

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Descriptive Statistics of the Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>Data (N = 46), n (%)</td>
</tr>
<tr>
<td>Gender</td>
<td>Male 26 (56.5) Female 20 (43.5)</td>
</tr>
<tr>
<td>Race</td>
<td>White 28 (60.9) Black 18 (39.1)</td>
</tr>
<tr>
<td>Age group, y</td>
<td>7–10 9 (19.6) 11–14 15 (32.6) 15–17 18 (39.1) 18–19 4 (8.7)</td>
</tr>
<tr>
<td>Comorbidities</td>
<td>Obesity 26 (56.5) Down syndrome 6 (13.0) Myelomeningocele 2 (4.4) Obesity, Asperger syndrome 1 (2.2) Obesity, autism 1 (2.2) Obesity, Chiari malformation 1 (2.2) Obesity, Down syndrome 1 (2.2) Obesity, Prader-Willi syndrome 1 (2.2) Allergic rhinitis 1 (2.2) Cerebral palsy 1 (2.2) Developmental delay, proxis 1 (2.2) Fragile X 1 (2.2) Muscular dystrophy 1 (2.2) Otopalatodigial syndrome 1 (2.2) Tourette’s syndrome, ADHD 1 (2.2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Clinical Data at Baseline (After Adenotonsillectomy) and at Optimal PAP Pressure Determined by Polysomnogram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Baseline</td>
</tr>
<tr>
<td>Patients spending &gt;10% of nights with ETCO₂ &gt;50 mm Hg (obstructive hypoventilation), n/N (%)</td>
<td>24/46 (52.2)</td>
</tr>
<tr>
<td>Percentage of nights at ETCO₂ &gt;50 mm Hg</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>SpO₂ nadir, %</td>
<td></td>
</tr>
<tr>
<td>ETCO₂, mm Hg</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>AHI events per h</td>
<td>Mean ± SD</td>
</tr>
</tbody>
</table>

ADHD indicates attention-deficit/hyperactivity disorder.

*Adherence was defined as >4 hours per night for ≥5 nights per week of PAP use and was able to be determined only for the 27 patients with meter readings.

b Data are based on comparisons of 44 patients with baseline polysomnogram and PAP data.

Data are from Wilcoxon signed-rank tests of significance of the difference between baseline polysomnogram and PAP mode.
ments in AHI, Spo₂ nadir, mean ETCO₂ level, and the presence of obstructive hypoventilation after PAP (P < .001).

After home PAP use, there was a significant decrease in nocturnal symptoms reported at the follow-up clinic visit (each P < .001). Significant improvements in daytime somnolence and school performance were also observed (each P < .001); improvement in hyperactivity was not statistically significant (P = .053).

PAP Adherence
The frequency of use and adherence could be calculated only for the 27 patients for whom meter readings were available. Of these 27 patients, 19 (70%) were determined to be adherent, using PAP a mean 73% (SD: 28%) of the week. Twenty-three patients (85%) used PAP >4 hours per night, with mean of 7.0 hours per night (SD: 2.1 hours per night) of use; the mean parent report for these 27 patients was 7.7 hours per night (SD: 1.7 hours per night). Figure 1 shows histograms of the number of patients by hours per night of PAP use based on parent report (Fig 1A) and meter readings (Fig 1B). As the Bland-Altman plot (Fig 2) illustrates, parental report showed good agreement with meter readings for 24 of the 27 patients having both parent and meter readings (agreement would not be considered “good” for the points above or below the 95% confidence limit). Overall, the limits of agreement (~1.91 and 3.30) were small enough to indicate that, whereas not perfect, parental report showed good agreement with meter readings. However, as shown in Table 4, parents of patients with meter readings of ≤4 hours per night of use overestimated the mean hours of PAP use per night by ~2 hours compared with the actual meter readings, whereas for patients with meter readings >4 hours per night, parents overestimated the mean hours of PAP use per night by only half an hour.

Neither the presence of subjective OSA symptoms, either alone or in total, nor the severity of OSA as defined by baseline AHI, Spo₂ nadir, mean ETCO₂ levels, or the presence of obstructive hypoventilation correlated with meter measures of hours of PAP use per night in

<p>| Table 3: Symptoms Reported and OSA Clinical Markers Before and After Home PAP Use |
|-----------------|----------------|----------------|-----------------|</p>
<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>With PAP</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients reporting symptoms, n/N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apnea</td>
<td>30/46 (65.2)</td>
<td>0/45 (0)</td>
<td>&lt; .001b</td>
</tr>
<tr>
<td>Enuresis</td>
<td>16/45 (35.6)</td>
<td>3/44 (6.8)</td>
<td>&lt; .001c</td>
</tr>
<tr>
<td>Snoring</td>
<td>41/46 (89.1)</td>
<td>0/45 (0)</td>
<td>&lt; .001b</td>
</tr>
<tr>
<td>Choking</td>
<td>35/46 (76.1)</td>
<td>0/45 (0)</td>
<td>&lt; .001b</td>
</tr>
<tr>
<td>Daytime somnolence</td>
<td>40/45 (88.9)</td>
<td>3/44 (6.8)</td>
<td>&lt; .001c</td>
</tr>
<tr>
<td>Hyperactivity or behavior problems</td>
<td>8/45 (17.8)</td>
<td>3/44 (6.8)</td>
<td>.053c</td>
</tr>
<tr>
<td>Deteriorating school problems</td>
<td>13/41 (31.7)</td>
<td>0/37 (0)</td>
<td>&lt; .001d</td>
</tr>
</tbody>
</table>

Symptomse
Mean ± SD           3.5 ± 1.2       0.1 ± 0.4        < .001f
Range               0–5            0–2              NA

NA indicates not applicable.

a Data are from Wilcoxon signed-rank tests of significance of the difference between pre-PAP and post-PAP measures reported.
b Data are based on comparisons of 45 patients with pre-PAP and post-PAP data.
c Data are based on comparisons of 44 patients with pre-PAP and post-PAP data.
d Data are based on comparisons of 36 patients with pre-PAP and post-PAP data.
e Data are the mean of the total of 5 symptoms: apnea, enuresis, snoring, choking, and daytime somnolence.
f Data are based on comparisons of 46 patients with pre-PAP and post-PAP data.
our patients. However, when the frequency of use was also taken into account, patients who were adherent had higher baseline AHI ($P/H_{11005}.022$) and showed greater improvement in AHI on PAP ($P/H_{11005}.015$) compared with nonadherent patients. The proportion of adherent patients did not differ significantly by age, gender, BMI, or race. The mode of PAP therapy (single versus bilevel pressure) was not significantly associated with adherence (Fisher’s exact test, $P/H_{11005}.145$), although 80% (16 of 20) of patients on single pressure were found to be adherent compared with 43% (3 of 7) of patients on bilevel pressure. For the 27 patients with meter readings, the average number of hours of PAP use per night, either by parental report ($P = .004$) or meter reading ($P = .005$), correlated significantly with the frequency of use (ie, the percentage of nights that PAP was used during the week).

DISCUSSION

Similar to previous studies in adults$^6,7$ and children$^8-12,20$ our study showed PAP to be associated with improvement in clinical symptoms and objective improvement in polysomnogram parameters. Although our study showed no significant change in reported hyperactivity or behavior problems, this may be because of the small number of patients reporting this symptom, because only 8 (18%) of 45 patients had hyperactivity or behavior problems before initiation of the PAP treatment. That 65% of patients had obesity as a comorbid condition deserves attention, given the growing problem of obesity in the United States$^{34}$ and growing interest in obesity as a risk factor for OSAS in children.$^{35}$

Contrary to earlier observations in children, which did not use objective meter readings,$^9,10$ but comparable to observations in adults,$^{17,32}$ 85% of our sample used PAP >4 hours per night based on clock-counter/meter reports (Table 4). Moreover, similar to findings in adults,$^{18}$ we found that mean hours of PAP use per night correlated positively with frequency of PAP use during the week. In contrast to 1 study showing regular PAP use in only 46% of adult patients,$^{19}$ 70% of our sample used PAP >4 hours per night for an average of 73% of the week. In a recent pediatric study, one third of the children dropped out before 6 months.$^{20}$ In contrast, we collected follow-up data for all of our patients after PAP was set up at home, with only 2 patients refusing PAP therapy. However, our study was a retrospective review of patients’ charts, whereas the Marcus et al study$^{20}$ used a prospective design. Although we asked patients about problems that they experienced with PAP use, their reasons for adherence to PAP therapy were not identified. It is possible that patients who were poorly motivated to use PAP at the outset did not keep their clinic appointments to have their initial polysomnogram or to have PAP started and, therefore, were not included in the chart review. Moreover, our sample consisted of school-aged children and adolescents with OSAS-associated comorbidities and persistence of OSA symptoms despite adenotonsillectomy. This persistence of OSA, in itself, may have motivated these patients to continue with the recommended PAP therapy. However, we can-

**TABLE 4** Comparison of Parent Report and Meter Readings of Mean Hours per Night of PAP Use, by Meter-Determined Adherence Groups (≤4 Hours per Night vs >4 Hours per Night)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Meter ≤4 h per Night (n = 4)</th>
<th>Meter &gt;4 h per Night (n = 23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual meter readings, h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>3.31 (1.03)</td>
<td>7.65 (1.48)</td>
</tr>
<tr>
<td>Range</td>
<td>2–5</td>
<td>5–10</td>
</tr>
<tr>
<td>Parental report, h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>5.25 (1.89)</td>
<td>8.13 (1.22)</td>
</tr>
<tr>
<td>Range</td>
<td>4–8</td>
<td>5–10</td>
</tr>
</tbody>
</table>
not attribute the persistence of OSAS as a cause of adherence to PAP because of the limitations of our retrospective study design.

Our results suggest that PAP adherence can be achieved among school-aged children and adolescents, a group deemed to be least compliant in earlier studies.\textsuperscript{9,10} This finding is notable given that adherence to medication among adult patients with chronic conditions has been reported to range from only 43\% to 78\% and drops dramatically after the first 6 months of therapy.\textsuperscript{36} Similarly, \textasciitilde 50\% of adolescents with chronic conditions do not comply with care recommendations.\textsuperscript{37} Among adult patients accepting home PAP, 12\% to 25\% can be expected to discontinue PAP use by 3 years.\textsuperscript{26} Long-term use of PAP by school-aged children and adolescents has not been reported previously; patients in our study had used PAP for a mean duration of 18.1 months at the time data were collected by chart review, ranging from 3 to 43 months of use among the 27 patients with meter readings.

Factors that influence adolescents’ adherence to chronic disease treatments have been studied with inconclusive findings.\textsuperscript{37} The severity of OSA as defined by baseline AHI, \textit{SpO2} nadir, and symptom scores was not significantly correlated with meter measures of hours of PAP use per night in our patients, similar to a recent report on adherence to PAP in children 2 to 16 years of age.\textsuperscript{20} As reported previously in adult studies, however, the mean number of hours of use per night alone may not be an accurate reflection of actual PAP use and adherence.\textsuperscript{18} Using a more restrictive definition of adherence, including frequency of nights used per week, as well as mean number of hours of PAP use per night, we found that baseline AHI and change in AHI on PAP differed significantly between adherent and nonadherent PAP users, similar to findings in adults.\textsuperscript{26,28} On the basis of our findings, we submit that a more conservative definition of adherence that includes both the number of hours and frequency of use would provide a more clinically meaningful definition of adherence than merely the mean number of hours per night used.

We found no significant associations between PAP adherence and gender, age, or BMI in our pediatric sample, similar to other reports in adults\textsuperscript{18,32} and in children.\textsuperscript{20} In addition, there was no significant difference between PAP mode (single versus bilevel pressure) and either PAP effectiveness or adherence among the 27 patients with meter readings, which previous studies also reported.\textsuperscript{20,38} Our sample size, however, may be too small to identify statistically significant associations between PAP adherence and these other variables (with increased risk of a type II error). The calculation was based on 27 patients with meter readings, with 20 patients on single and only 7 patients on bilevel pressures, and the patients were not randomly assigned to treatment group. Thus, a prospective randomized trial to compare adherence using single and bilevel PAP in a larger sample of school-aged children and adolescents is recommended.

Studies in adults have suggested that intense patient education is associated with improved adherence.\textsuperscript{17,25} Our patients and their parents were all educated about OSAS and PAP by a pediatric, board-certified sleep specialist. Information alone, however, may not be enough to promote behavioral change in adolescents. Anxiety has been found to be related to nonadherence to PAP therapy in adults.\textsuperscript{39} Whether anxiety is associated with nonadherence to PAP therapy in adolescents is unknown. For chronically ill adolescents, attitudes, as well as the personal meaning and significance of an illness and its treatment, have been cited as an important factor affecting adherence.\textsuperscript{37} In addition, the importance of a good relationship between the adolescent patient and health care staff cannot be overemphasized.\textsuperscript{35} We believe that the relatively high level of adherence to PAP in our study may have been related to the continuity in care and follow-up of our patients, but an experimental design would be necessary to test this hypothesis as well. How their adherence to PAP may change over time also remains to be seen.

Much has been written about the differences in adolescents’ cognitive development as they go through early (11–14 years), middle (15–17 years), and late (17–21) adolescence.\textsuperscript{23} More than 70\% of our patients were in the early and midadolescent age groups, when formal operative thinking has yet to develop and children still have limited ability to consider long-range health risks.\textsuperscript{40} Nevertheless, we observed no significant differences in adherence among the 4 age groups, the youngest age group being \textless 11 years of age. Moreover, only 2 (4\%) of the 46 patients refused PAP; 1 was 11 years old (early adolescence) and the other was 19 years old (late adolescence). This low level of refusal to use PAP is similar to the report by McArdle et al in adults,\textsuperscript{29} but less than the 8\% to 15\% reported by others.\textsuperscript{13,14} Janson et al\textsuperscript{41} identified the lack of subjective effect of treatment as a common reason for nonadherence among his adult patients, and lack of perceived benefit might have been the reason for nonuse of PAP in the 2 patients who refused PAP, because both of these patients had very mild OSA at baseline (AHI of 1.5 and 6.4 events per hour).

Similar to recent studies,\textsuperscript{20,21} we only included children and adolescents who had undergone otolaryngology evaluation and adenotonsillectomy, a potential confounding variable and the treatment of choice in the pediatric population.\textsuperscript{4} If present after adenotonsillectomy, OSAS must be viewed as a chronic illness and PAP as long-term treatment. Future research might be designed to evaluate differences in adherence as a function of school-age and adolescent cognitive development using a prospective study design. Better understanding of adolescents’ perceptions of OSAS and how their percep-
tions affect adherence to PAP therapy are warranted. Moreover, whether adolescents who are adherent with PAP therapy continue to be adherent as adults remains to be seen.

This study has limitations inherent to its small sample size and retrospective design. Intensive support and behavior intervention have been recommended to improve PAP adherence. However, although we think our educational intervention and regular follow-up of patients constituted an effective and comprehensive support mechanism, we cannot infer that our intervention “caused” adherence to PAP resulting in improved outcomes, because we analyzed existing data, and the intervention was offered to all of the patients. Also, whereas parent-reported hours of use per night showed good agreement with meter readings in this study (but for a few points outside the confidence bounds in Fig 2), studies have shown that self-reported use in adults overestimated PAP use by an average of 1.0 hours when compared with the more objective meter readings.

A recent study of PAP adherence in children also showed parental assessment overestimated actual use, and we found that parents of patients with meter readings of ≤4 hours per night overestimated the mean hours of PAP use per night by ~2 hours (Table 4). An important difference between the study by Marcus et al. and our study is that we reported better adherence among our school-aged children and adolescents, with more patients in our study registering >4 hours of PAP use per night on the meter than registering ≤4 hours per night. But with 27 patients in all (and only 4 in the group registering ≤4 hours per night on the meter), our finding of good agreement between parent-reported use and meter readings may be a “false-negative” finding and may be attributed to increased β error (low power to detect a statistically significant difference). Regardless, the contrast in findings between the Marcus et al. study and ours speaks to the necessity of additional research using an experimental design to better assess the external validity of findings from each of the 2 studies. In addition, in this retrospective study, not all of the patients were blinded to the presence of a clock-counter meter. It remains to be seen whether having knowledge of the presence of a clock-counter meter correlates with better PAP adherence.

In addition, whereas adherence had been defined in terms of hours and frequency of use, sleep patterns and habits also may be important, given reports that OSAS is predominantly a rapid eye movement phenomenon in children. Because children tend to consolidate rapid eye movement sleep toward the latter half of the night and the early morning, when most PAP sessions might end, consideration as to what constitutes regular use might need to include not only the number of hours used but, specifically, during which hours of sleep it is used.

CONCLUSIONS

This study shows that most of the school-aged children and adolescents who received a dedicated education and follow-up program at our institution were adherent to PAP therapy. Patients with a higher baseline AHI and a greater change in AHI on PAP were more likely to be adherent. Adherence was not associated with gender, race, age, or BMI in our sample of pediatric patients. Identification of other factors that may be associated with PAP adherence by school-aged children and adolescent patients requires additional study.

ACKNOWLEDGMENT

We thank Yan Yan, MD, PhD, for constructing the Bland-Altman plot and providing other statistical analytical support.

REFERENCES


Cerebrospinal Fluid Xanthochromia in Newborns Is Related to Maternal Labor Before Delivery

Lise E. Nigrovic, MD, MPH, Michelle Trivedi, MS, Jonathan A. Edlow, MD, Mark I. Neuman, MD, MPH

ABSTRACT

OBJECTIVE. The purpose of this work was to investigate whether xanthochromia in newborns is related to maternal labor before delivery.

METHODS. We reviewed the medical charts of all of the infants ≤ 30 days of age who had a lumbar puncture performed in a single pediatric emergency department between 2003 and 2005. Xanthochromia was detected by the hospital laboratory using the qualitative visual inspection method. We used logistic regression to determine the relationship between maternal labor before birth and the presence of cerebrospinal fluid xanthochromia, adjusting for factors known to be associated with xanthochromia.

RESULTS. Of the 478 newborns who had a lumbar puncture performed during the study period, 134 (28%) had xanthochromia. Of the 449 infants with delivery method recorded in the medical chart, 332 (74%) were born via vaginal delivery, 24 (5%) via cesarean section after maternal labor, and 93 (21%) via cesarean section without maternal labor. After excluding patients with hyperbilirubinemia (total bilirubin ≥ 15 mg/dL) and adjusting for factors known to be associated with xanthochromia (cerebrospinal fluid red blood cells ≥ 20 000 cells per mL and cerebrospinal fluid protein ≥ 150 mg/dL), infants born after maternal labor had a higher rate of cerebrospinal fluid xanthochromia than infants born without any labor.

CONCLUSIONS. Xanthochromia is a common finding in the cerebrospinal fluid of newborns and is associated with maternal labor preceding delivery.
Xanthochromia is the yellow discoloration of cerebrospinal fluid (CSF) that results from the presence of pigmented products of hemoglobin catabolism. When red blood cells enter the subarachnoid space, they are gradually lysed, and the released hemoglobin is then metabolized to the pigmented molecules oxyhemoglobin, methemoglobin, and bilirubin.1,2 This enzyme-dependent breakdown process begins after several hours1 but may take up to 12 hours after the onset of bleeding to develop.4,5 Xanthochromia is detected by either visual inspection or by spectrophotometry of the CSF.1 Xanthochromia has been observed in patients with an elevated serum bilirubin (≥15 mg/dL),6 elevated CSF protein (≥150 mg/dL),7 and in infants in whom the lumbar puncture (LP) is traumatic or yields a grossly bloody result.5,8 In an in vitro study, 70% of CSF specimens with ≥20,000 red blood cells (RBCs) per mL had xanthochromia when samples were analyzed immediately compared with 100% if analysis was delayed by 1 hour.8

In adults presenting with acute-onset headaches, xanthochromia has been used as a marker for subarachnoid hemorrhage;4 a diagnosis that is an unusual consideration for young infants.9 The presence of xanthochromia in the CSF has also been observed among neonates with herpes simplex virus (HSV) meningoencephalitis.10 Although based on a single case series of Japanese children, the presence of xanthochromia alone often leads to the consideration of HSV infection in the youngest infants. The clinical significance of xanthochromia in newborn CSF is not well understood.

Maternal labor preceding birth has been associated with occult intracranial hemorrhage in asymptomatic newborns.11 We hypothesize that RBCs that enter the CSF during labor may later be catabolized, thus resulting in xanthochromia. To investigate the relationship between xanthochromia in the CSF of infants and maternal labor, we reviewed the medical charts of ~500 infants who underwent a LP in the emergency department during the first month of life.

METHODS

Study Design
We performed a retrospective cohort study of all of the infants (≥30 days of age) who had an LP performed in the emergency department of the study institution during a 3-year period (January 2003 to December 2005). Eligible patients were identified by query of a laboratory database and were included if they had CSF obtained in the emergency department. We included all of the patients with an available CSF cell count. As part of standard laboratory protocol, all of these patients had CSF color examined.

The emergency department and inpatient medical charts of all of the study patients were retrospectively reviewed to determine patient demographics, mode of infant delivery (vaginal delivery, cesarean section after maternal labor, or cesarean section without maternal labor), clinical indication for LP, and patient disposition (discharged from the emergency department or admitted to the hospital). Laboratory databases were reviewed for serum bilirubin and CSF results (RBCs, white blood cells, protein, glucose, and color). Xanthochromia was determined at the time of specimen collection by laboratory personnel using centrifuged CSF specimen and qualitative visual inspection methods.1 Data were abstracted about neuroimaging and HSV testing, when either was performed.

Statistical Analysis
We performed univariate testing with χ² analysis. We performed multiple logistic regression to determine the association between maternal labor and CSF xanthochromia. We excluded all of the patients with serum bilirubin levels of ≥15 mg/dL, because xanthochromia may be because of an elevated serum bilirubin, independent of other factors.5,7 We did not exclude patients with intracranial hemorrhage identified on neuroimaging performed after the LP.

In the multivariate analysis, we adjusted both for patient age (in days) and for factors known to be associated with xanthochromia (elevated CSF RBC and CSF protein).7,8 We dichotomized CSF RBCs (≥20,000 cells per mL)8 and CSF protein (≥150 mg/dL)7 based on published thresholds associated with xanthochromia. Only patients with complete data on all of the candidate predictors were included in the multivariable analysis.

<table>
<thead>
<tr>
<th>TABLE 1 Patient (N = 478) Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Age, median (interquartile range), d</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
</tr>
<tr>
<td>Race, n (%)</td>
</tr>
<tr>
<td>White</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Delivery method, n (%)</td>
</tr>
<tr>
<td>Vaginal delivery</td>
</tr>
<tr>
<td>Cesarean section after any labor</td>
</tr>
<tr>
<td>Cesarean section with no labor</td>
</tr>
<tr>
<td>Clinical indication for LP, n (%)</td>
</tr>
<tr>
<td>Neonatal fever</td>
</tr>
<tr>
<td>Seizure</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Serum bilirubin level of ≥15 mg/dL, n (%)</td>
</tr>
<tr>
<td>CSF white blood cell count, median (interquartile range), cells per mL</td>
</tr>
<tr>
<td>CSF RBC count, median (interquartile range), cells per mL</td>
</tr>
<tr>
<td>CSF glucose, median (interquartile range), mg/dL</td>
</tr>
<tr>
<td>CSF protein, median (interquartile range), mg/dL</td>
</tr>
<tr>
<td>Xanthochromia, n (%)</td>
</tr>
<tr>
<td>Admitted to the hospital, n (%)</td>
</tr>
</tbody>
</table>

* Serum bilirubin was obtained from 92 infants, of whom, 15 had a bilirubin level of ≥15 mg/dL.
For all of the analyses, we used SPSS (SPSS Inc, Chicago, IL). The study was approved by the institutional review board of the study institution.

RESULTS
We identified 478 infants ≤30 days of age who had a CSF obtained for analysis from an LP performed in the emergency department during the study period. No patient had >1 LP performed within the first 30 days of life during the study period.

Table 1 provides the clinical and laboratory characteristics of the patients included in the study. LPs were obtained for the following clinical reasons: 89% for fever in a neonate, 9% for seizure, and 2% for other reasons. Of the 449 study patients with birth method recorded in the medical chart, 332 (74%) were born after vaginal delivery, 24 (5%) by cesarean section after maternal labor, and 93 (21%) by cesarean section without maternal labor. A total of 134 infants (28%) had xanthochromia noted in the CSF.

In the univariate analysis, CSF xanthochromia varied by birth method: it was most common in infants born by vaginal delivery and least common in infants born by cesarean section without maternal labor. A total of 134 infants (28%) had xanthochromia noted in the CSF. After adjusting for patient age, elevated CSF protein (≥150 mg/dL), and bloody CSF (CSF RBCs ≥20 000 cells per mL), we found that maternal history of labor before infant delivery was independently associated with the presence of xanthochromia (adjusted odds ratio: 2.4; 95% confidence interval: 1.2–5.0; Table 3).

Of the 478 study patients, 71 patients (15%) had neuroimaging performed (computed tomography, MRI, or ultrasound examinations of the head). Four of these patients had intracranial hemorrhage identified (3 with subdural hematomas and 1 with intraparenchymal blood). Xanthochromia was observed in the 1 infant with an intraparenchymal bleed; none of the 3 infants with subdural hematomas had xanthochromia noted in the CSF.

Sixty-one (13%) of patients had HSV polymerase chain reaction (PCR) sent from the CSF, of which, 1 infant tested positive. This infant was 11 days of age and presented with a vesicular rash suggestive of disseminated herpes infection. The CSF from this infant was xanthochromic and had a CSF RBC of 385 cells per mL.

DISCUSSION
In reviewing the CSF results for all of the infants undergoing LP in the first month of life, we found CSF xanthochromia to be relatively common among neonates (28% of all infants ≤30 days of life). First, we observed an inverse relationship between patient age and the presence of xanthochromia. Second, we found an independent association between history of maternal labor before delivery and CSF xanthochromia after excluding those infants with hyperbilirubinemia and adjusting for other factors reported to be associated with xanthochromia, including grossly bloody CSF and elevated CSF protein.

We hypothesized that maternal labor may cause sub-
clinical intracranial hemorrhage in some neonates. The hemoglobin from these RBCs that enter the subarachnoid spaces is later to be enzymatically degraded to the pigmented products (oxyhemoglobin, methemoglobin, and bilirubin), which may persist for several weeks. For those infants who later undergo LP, the resulting CSF may be xanthochromic.

Other investigators have demonstrated occult intracranial hemorrhage in asymptomatic newborns. Whitby et al performed cranial MRIs on 111 normal asymptomatic newborns. Nine infants (8% of study patients) had subdural hemorrhages identified all by MRI performed at 48 hours of life. All of the hemorrhages had spontaneously resolved by the time of a repeat MRI performed at 4 weeks of life. The incidence of xanthochromia in infants with subdural hematomas is unknown; however, none of the 3 infants with subdural hematomas in our study had xanthochromia.

Investigators have previously identified an association between mode of infant delivery and intracranial hemorrhage. In a birth cohort of >500,000 newborns, Towne et al found that forceful delivery (either use of forceps or vacuum extraction) was associated with an increased risk of clinically significant intracranial hemorrhage. In a second study, Looney et al performed MRIs on 97 neonates between 1 and 5 weeks of life. Seventeen neonates (18% of study patients) had intracranial hemorrhages identified; all of these infants were born vaginally. Although some of these infants likely had an LP performed in the neonatal period, neither study reported CSF findings.

CSF xanthochromia has been described in children with HSV encephalitis. Neonatal HSV infection is an important clinical concern among febrile neonates and may result in severe neurologic sequelae. In our study, only 13% of patients had CSF herpes PCR tests sent. The only infant who had a positive HSV PCR did have CSF xanthochromia but also had a classic whole-body vesicular rash. Given the rarity of HSV infections, our study was not powered to determine the sensitivity of CSF xanthochromia for HSV encephalitis. However, in almost 500 infant LPs, the presence of xanthochromia did not identify a single case of unsuspected HSV infection. Given the frequency with which CSF xanthochromia is observed in healthy neonates (28%) and the rarity of herpes meningoencephalitis in well-appearing infants without seizures or vesicular rash, we would argue that CSF xanthochromia alone should not prompt consideration of neonatal HSV infection. With that said,

---

**TABLE 3**  Multivariable Logistic Regression Analysis for the Presence of CSF Xanthochromia

<table>
<thead>
<tr>
<th>Variable</th>
<th>CSF Xanthochromia (n = 123 of 463 LPs), Adjusted Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any labor</td>
<td>2.4 (1.2–5.0)</td>
</tr>
<tr>
<td>Age, d</td>
<td>0.9 (0.9–0.9)</td>
</tr>
<tr>
<td>CSF protein ≥150 mg/dL</td>
<td>2.6 (1.1–6.0)</td>
</tr>
<tr>
<td>CSF RBCs ≥20,000 cells per mL</td>
<td>1.7 (0.8–3.8)</td>
</tr>
</tbody>
</table>

---

![FIGURE 1](https://example.com/figure1.png) Percentage of xanthochromia based on age of patient (days) at time of lumbar puncture.
one must always have a low suspicion for testing and
treatment for HSV among young febrile infants (<1
month of age), particularly if CSF pleocytosis is found.

Our study has the following limitations. First, our
study was retrospective, and birth method was collected
by medical chart review. However, only 6% of all of the
charts reviewed were lacking information regarding
birth method. In addition, for the 94% of patients who
were admitted to the hospital, we noted perfect correla-
tion between the emergency department and inpatient
records for birth method. Second, CSF xanthochromia
was determined by using visual inspection rather than
spectrometry. Although visual inspection may be less
sensitive test than spectrometry, this technique is
used by almost all institutions (99.7% of US clinical
laboratories in a recent survey). In addition, spectro-
photometry has only moderate-to-low specificity, and
the literature suggests that the discrepancy between vi-
sual detection methods and spectrophotometry is not as
great as previous studies have claimed. Third, we do not
know the length of time between the physician
performing the LP and the analysis of the CSF by the
clinical laboratory. Previous investigators have demon-
strated that xanthochromia increases with a delay in
specimen processing. However, after discussion with
technicians in the clinical laboratory, these delays should
have been minimal and certainly would be unrelated to
the mode of infant delivery. Furthermore, only 19% of
patients had serum bilirubin obtained at the time of their
emergency department evaluation. However, patients
with a serum bilirubin level of ≥15 mg/dL would likely
appear jaundiced, and most clinicians would obtain a
bilirubin measurement. Last, only a minority of study
patients had neuroimaging performed. Therefore, we
could not assess the correlation between CSF xantho-
chromia and radiographic intracranial hemorrhage in
our study population. With that said, patients with clin-
ical presentations concerning for significant hemorrhage
would have had head imaging performed.

CONCLUSIONS
We found CSF xanthochromia to be common among
young infants, and it is associated with a history of
maternal labor. We hypothesize that maternal labor may
be associated with subclinical hemorrhage in some new-
borns, with resulting CSF xanthochromia, which may
persist throughout the first month of life. Therefore, the
finding of CSF xanthochromia alone in an infant ≥30
days of age should not, in and of itself, prompt consid-
eration of significant intracranial hemorrhage.

ACKNOWLEDGMENTS
This work was supported by National Research Service
Awards (grant T32 HD40128-01 Research Training in Pe-
diatric Emergency Medicine and grant T32 HD40304-04
Research Training in Pediatrics [to Dr Nigrovic]).

REFERENCES
1. Edlow JA, Bruner KS, Horowitz GL. Xanthochromia. Arch
2. Shah KH, Edlow JA. Distinguishing traumatic lumbar puncture
from true subarachnoid hemorrhage. J Emerg Med. 2002;23:
67–74
Livingstone Ltd; 1956.
4. Dougherty JM, Roth RM. Cerebrospinal fluid. Emerg Med Clin
5. Suarez JL, Tarr RW, Selman WR. Aneurysmal subarachnoid
6. Fishman RA. Cerebrospinal Fluid in Diseases of the Nervous System.
7. Seehusen DA, Reeves MM, Fornin DA. Cerebrospinal fluid
8. Graves P, Sidman R. Xanthochromia is not pathognomonic for
9. Simbrunner J, Riccabona M. Imaging of the neonatal CNS. Eur
J Radiol. 2006;60:133–151.
10. Kamei S, Takasu T, Otani S, Mochizuki Y. Cerebrospinal fluid
findings in 108 Japanese cases of herpes simplex encephalitis
history of subdural haemorrhages in babies and relation to
12. SPSS for Windows [computer program]. Version 14.0.2. Chi-
cago, IL: SPSS Inc; 2005.
13. Towner D, Castro MA, Eby-Wilkens E, Gilbert WM. Effect of
mode of delivery in nulliparous women on neonatal intracra-
in asymptomatic neonates: prevalence on MR images and
relationship to obstetric and neonatal risk factors. Radiol-
15. Fidler KJ, Pierce CM, Cubitt WD, Novelli V, Peters MJ. Could
neonatal disseminated herpes simplex virus infections be
reliably detect cerebrospinal fluid xanthochromia. Stroke. 2005;
36:1295–1297.
17. Petzold A, Keir G, Sharpe LT. Spectrophotometry for xen-
18. Perry JJ, Sivilitto ML, Stiell IG, et al. Should spectrophotome-
try be used to identify xanthochromia in the cerebrospinal
fluid of alert patients suspected of having subarachnoid hem-
inspection versus spectrophotometry in detecting bilirubin in
cerebrospinal fluid. J Neurol Neurosurg Psychiatry. 2005;76:
1452–1454.
Is There Differential Retention of Children With Special Health Care Needs in the State Children’s Health Insurance Program?

Tamarie Macon, BS*, Jane E. Miller, PhD*,b, Dorothy Gaboda, PhD*, Theresa Simpson, BS*, Joel C. Cantor, ScD*c

*aInstitute for Health, Health Care Policy, and Aging Research, bE. J. Bloustein School of Planning and Public Policy, and cCenter for State Health Policy, Rutgers University, New Brunswick, New Jersey

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The purpose of this work was to determine whether children with special health care needs in New Jersey’s State Children’s Health Insurance Program are less likely to become uninsured than children without special health care needs.

PATIENTS AND METHODS. We used the 2003 New Jersey FamilyCare Supplement to the New Jersey Family Health Survey. Children were randomly selected from the universe of children enrolled in New Jersey FamilyCare as of May 2002, and their families were surveyed during 2003 (N = 675). The Children With Special Health Care Needs Screener was used to identify 5 types of special health care needs. We estimated multinomial logistic regression models of final enrollment status according to the presence of ≥1 special health care need, controlling for demographic characteristics.

RESULTS. Roughly 1 of every 5 children in New Jersey FamilyCare had ≥1 special health care need. Older children and boys had greater odds of having special health care needs than others. Children with special health care needs had only one fourth the odds of becoming disenrolled and uninsured compared with children without special health care needs, even when controlling for age, gender, race/ethnicity, and insurance plan level. There was no difference in likelihood of finding other health insurance according to children with special health care needs status.

CONCLUSIONS. Children with special health care needs were more likely than children without such needs to be covered by health insurance at the time of the survey, either by retaining State Children’s Health Insurance Program coverage or by finding other insurance. The higher retention of children with special health care needs in New Jersey FamilyCare is good news for families of these children and their advocates. However, higher health care costs for these children should be considered in federal and state budget planning for the State Children’s Health Insurance Program.
A DECADE AGO, the US government enacted one of the most successful pieces of legislation to reduce uninsurance among children, the State Children’s Health Insurance Program (SCHIP), which was created in 1997 under the Social Security Act to provide insurance to children from poor families who do not qualify for Medicaid.1 States welcomed SCHIP from the beginning, with >6 million children insured by the program as of 2005.2 Although rates of enrollment are useful to determine overall coverage rates, disenrollment patterns are equally critical because they capture which children are staying in the program. As Congress debates the reauthorization in 2007, it is important to know the health status and other characteristics of the SCHIP-insured population that might affect costs.3

Ensuring continuity of coverage in publicly funded programs such as SCHIP is critical for all children, but especially for those with special needs, such as asthma or diabetes, because these children require more and more frequent health care services. Children with special health care needs (CSHCN) are a special population defined by the Maternal and Child Health Bureau as “those who have or are at increased risk for a chronic, physical, developmental, behavioral, or emotional condition and who also require health and related services of a type or amount beyond that required by children generally.”4 A recent study of 5 SCHIPs found that children with special needs were more likely than children without special needs to use the emergency department, mental health care, specialty care, and acute care.5

According to the 2001 National Survey on Children With Special Health Care Needs, ~9.5 million children in the United States have special health care needs (SHCNs).6 The CSHCN screener is a noncategorical, outcome-based approach to identifying children with special needs.7 Rather than asking about specific childhood chronic conditions (which overlooks many of the less common chronic conditions), this approach asks parents about their children’s need for or use of health care services beyond those typical of healthy children. Consequently, the measure captures children who have rare conditions or special needs, such as cerebral palsy or cystic fibrosis, as well as those afflicted by common chronic conditions, such as asthma or diabetes. This attribute is of special import for policy-makers and researchers who need to monitor the quality of care for all CSHCN who enroll in public health insurance programs.

CSHCN are defined as having more chronic health conditions than the general population of children; therefore, adverse differential selection and retention are of concern in terms of determining adequate funding levels for the SCHIP. Those selection processes imply that sicker people are more likely to enroll and remain enrolled than healthy people, causing the overall health of those in the SCHIP to be worse than those in the SCHIP-eligible population. It is important to note that differential selection and retention concern parents’ choices and actions related to SCHIP enrollment and renewal, not selectivity into SCHIP based on eligibility criteria for the program. Title XXI of the Social Security Act specifically forbids such “cherry-picking” in that states “may not deny eligibility based on a child having a preexisting medical condition.”8

Differential selection has been suggested by evidence from SCHIP initiatives in other states. An analysis of New York, Florida, and Kansas found that CSHCN make up between 17% and 25% of the total enrolled population,9 higher than the estimates of 13% CSHCN in the general population based on the National Survey on Children With Special Health Care Needs,6 and 17% among low-income uninsured children from a recent study by Newacheck et al.10 In addition, a study of Florida’s SCHIP using the Questionnaire for Identifying Children with Chronic Conditions found that enrolled children were twice as likely to have chronic conditions as children in the income-restricted national subsample.11 Because CSHCN are more concentrated in the SCHIP than in the SCHIP-eligible population, it seems that parents selectively choose to enroll their child based on health status. Moreover, a study of parental attitudes about the need for health insurance showed that parents of CSHCN were least likely to report that their child does not need insurance (2.8% vs 7.1% for non-CSHCN).12 An analysis of the Healthy Kids program in Florida showed that children with physical health needs were less likely to disenroll from the program and more likely to reenroll.13 Taken together, these studies suggest that parents of children with special needs recognize the greater incentive to keep their children enrolled in SCHIP. This study extends previous analyses of differential retention of CSHCN by contrasting 3 possible health insurance outcomes, controlling for demographic characteristics and SCHIP plan (which controls for family income) level: (1) remaining enrolled in the SCHIP, (2) disenrolling from the SCHIP but finding other health insurance, and (3) becoming uninsured on disenrollment. By investigating transitions out of the SCHIP according to SCHCN status, we complement studies that compare static (snapshot) CSHCN prevalence between SCHIP enrollees and children who are SCHIP-eligible but not enrolled.

METHODS
Data Source and Study Sample
We used data from the 2003 New Jersey FamilyCare Supplement (NJFCS), a telephone survey of families with children enrolled in New Jersey’s SCHIP as of May 2002. The NJFCS collected information about health status, access to care, insurance coverage, and health care use, as well as experiences with enrollment, renewal, disenrollment, and satisfaction with New Jersey
FamilyCare (NJFC; New Jersey’s SCHIP initiative) and the CSHCN screener (see below).14 The survey protocol and questionnaire were approved by the human subjects review board at the researchers’ university.

The NJFCS was conducted between May and September 2003, and the respondents were the adults most knowledgeable about the child in the program, who were asked to grant oral consent for the study at the start of the telephone interview. Families were randomly selected to participate if ≥1 of their children had been enrolled in the NJFC program in the previous year. To ensure adequate representation of both enrolled and disenrolled children, the sample was stratified according to enrollment status as of January 2003 based on administrative records. Children who were still enrolled and the majority of disenrolled children had been in the program for ≥1 year; disenrolled children left the plan between 1 and 12 months before their interview. The sample was also stratified by SCHIP plan level (defined below) and whether parents were also enrolled in NJFC, yielding 10 strata from which children were selected (see Appendix for a diagram of the strata).

A total of 679 families participated in the study, yielding an overall response rate of 52%. The response rate was ~10 percentage points lower among those who had disenrolled than among those who remained enrolled and ~5 percentage points lower among those without parents enrolled. Comparison against administrative records of all of the children enrolled in NJFC between 2000 and 2002 show that girls, non-Hispanic black children, and those <5 years old were underrepresented among survey respondents.

Children were chosen at random from among the universe of case subjects enrolled in NJFC as of May 2002 as the “index” children. If 2 children were chosen from the same family, the first chosen was selected as the index child before the telephone interview. Children who disenrolled and were no longer qualified for the NJFC program because of income or age requirements or who were missing data on final enrollment status were excluded from this analysis (n = 4), resulting in a final analytic sample of 675 index children. All of the estimates presented are weighted to reflect the universe of children enrolled in NJFC as of May 31, 2002, taking into account differential sampling probabilities for the strata shown in the Appendix.14

Measures

Identification of CSHCN

Children were identified as having special needs according to the CSHCN screener developed by the Child and Adolescent Health Measurement Initiative of the Foundation for Accountability.7,15 Based on the Maternal and Child Health Bureau’s definition of CSHCN, the screener includes the following questions to identify children experiencing health-related consequences that have lasted or are expected to last ≥12 months: (1) Do any of your children under 18 currently need or use medicine prescribed by a doctor (other than vitamins) because of any medical, behavioral, or other health condition? (2) Do any of your children under 18 need or use more medical care, mental health, or educational services than is usual for most children of the same age because of any medical, behavioral, or other health condition? (3) Are any of your children under 18 limited or prevented in any way in their ability to do the things most children of the same age can do because of any medical, behavioral, or other health condition? (4) Do any of your children under 18 have any kind of emotional, developmental, or behavioral problem for which he or she needs or gets treatment or counseling? To be classified as having a SHCN, the child must have been classified “yes” for ≥1 of the 5 questions about the different areas of need and the associated follow-up question on the condition lasting ≥12 months.

Final Enrollment Status in NJFC

Final enrollment status was based on survey responses verified against NJFC administrative records. Initial classification was determined from survey questions on whether the index child remained enrolled in the SCHIP at the time of the survey in 2003 and, if not, the presence of other health insurance, and family income (to assess continued eligibility). Comparison of the survey-based classification against NJFC administrative records revealed a few case subjects with short gaps in enrollment who were reclassified as “enrolled”; 18 case subjects enrolled in Medicaid who had reported themselves “not in SCHIP” were also classified as enrolled. These steps resulted in final enrollment status classified into 1 of 3 categories: (1) currently enrolled in the NJFC or Medicaid; (2) disenrolled from NJFC but still insured (either through parents’ employment or other insurance); and (3) disenrolled from NJFC, uninsured, but still eligible for the program. Two children who had disenrolled and were no longer eligible because of age or income requirements were excluded.

SCHIP Plan Level

The NJFC program has 1 of the highest income eligibility levels of all states, covering otherwise uninsured children from families with incomes ≥350% of the federal poverty level (FPL). The NJFC program is divided into 4 plan levels labeled A through D. Plan A is a Medicaid expansion program, which covers children in families up to 133% of the FPL, plans B and C cover those with family incomes between 133% and 200% of the FPL, and plan D provides coverage to children with family incomes ≥200% of the FPL.
incomes between 201% and 350% of the FPL. In 2003, 350% of the FPL was approximately $65,300 for a family of 2 adults and 2 children.16 Plans C and D involve cost sharing in the form of monthly premiums and copayments on a sliding scale based on family income. The survey combined plans B and C for sampling purposes because of the narrow range of incomes covered in each plan.

**Health Measures**
To validate the CSHCN screener, we compared it against 2 other health measures from the survey: parent’s rating of the child’s health (classified “excellent,” “very good,” “good,” “fair,” and “poor”) and an asthma indicator (yes or no).

**Control Variables**
To control for potential confounding, several demographic characteristics were also considered, including age, gender, and race/ethnicity coded as shown in Table 1.

### Data Analyses
$\chi^2$ tests were conducted to determine associations between demographic and plan factors and the prevalence of SHCN. Multinomial logistic regression models were used to assess differential retention of CSHCN by comparing the odds of being (1) disenrolled and insured and (2) disenrolled and uninsured compared with children who were still enrolled in NJFC, controlling for demographic factors. All of the statistics were weighted to the state level population of children enrolled in NJFC as of May 31, 2002, using sampling weights provided for the NJFC survey.14

### RESULTS
Table 1 presents the demographic and health characteristics of the study sample, as well as the prevalence of SHCN according to those characteristics. Prevalence of SHCN was higher among older children and boys ($P < .01$). Prevalence of SHCN among children aged 6 to 12

---

**TABLE 1**  
**Demographic and Health Characteristics of Children in the 2003 NJFC Sample**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Sample Composition</th>
<th>Prevalence of CSHCN$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unweighted N</td>
<td>Weighted % of Sample$^b$</td>
</tr>
<tr>
<td>Whole sample</td>
<td>675</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Demographic characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0- to 5-y-olds</td>
<td>113</td>
<td>15.6</td>
</tr>
<tr>
<td>6- to 12-y-olds</td>
<td>332</td>
<td>44.2</td>
</tr>
<tr>
<td>13- to 18-y-olds</td>
<td>230</td>
<td>40.1</td>
</tr>
<tr>
<td>Gender of child</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girl</td>
<td>321</td>
<td>40.2</td>
</tr>
<tr>
<td>Boy</td>
<td>354</td>
<td>59.8</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>290</td>
<td>39.9</td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>104</td>
<td>15.8</td>
</tr>
<tr>
<td>Hispanic</td>
<td>220</td>
<td>36.5</td>
</tr>
<tr>
<td>Other race</td>
<td>61</td>
<td>7.7</td>
</tr>
<tr>
<td><strong>SCHIP characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NJFC plan level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plan A (≤133% of FPL)</td>
<td>185</td>
<td>37.5</td>
</tr>
<tr>
<td>Plans B and C (133%–200% of FPL)</td>
<td>244</td>
<td>45.0</td>
</tr>
<tr>
<td>Plan D (201%–350% of FPL)</td>
<td>246</td>
<td>17.5</td>
</tr>
<tr>
<td><strong>Health status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent-rated health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent</td>
<td>260</td>
<td>39.3</td>
</tr>
<tr>
<td>Very good</td>
<td>169</td>
<td>23.8</td>
</tr>
<tr>
<td>Good</td>
<td>194</td>
<td>29.3</td>
</tr>
<tr>
<td>Fair</td>
<td>38</td>
<td>6.4</td>
</tr>
<tr>
<td>Poor</td>
<td>14</td>
<td>1.2</td>
</tr>
<tr>
<td>Asthma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>574</td>
<td>86.3</td>
</tr>
<tr>
<td>Yes</td>
<td>101</td>
<td>13.7</td>
</tr>
<tr>
<td><strong>Final enrollment status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrolled in NJFC or Medicaid</td>
<td>444</td>
<td>87.5</td>
</tr>
<tr>
<td>Disenrolled but insured</td>
<td>145</td>
<td>6.6</td>
</tr>
<tr>
<td>Disenrolled and uninsured</td>
<td>86</td>
<td>5.9</td>
</tr>
</tbody>
</table>

---

Data were weighted to the universe of all of the children enrolled in the NJFC program as of May 2002.

$^a$ Data show the percentage of children with ≥1 SHCN as identified by CSHCN screener.

$^b$ P < .01.

$^c$ P < .001.
and 13 to 18 years was twice as high as among those aged 1 to 5 years (25% and 20% vs 9%). Boys had \( \sim 1.5 \) times the odds of having special needs as girls (24% and 16%, respectively). The prevalence of children with health care needs was roughly equal in the different SCHIP plans (across incomes \( \leq 350\% \) of the FPL), with \( \sim 1 \) in 5 children at each plan level having \( \geq 1 \) SHCN.

Overall, the prevalence of SHCN was 21% of the SCHIP sample, and the mean number of SHCN was 0.45. The most common type of SHCN was medication use (15.1% of enrolled children), followed by greater health services than most children (10.3%), activity limitations (8.1%), emotional problems (7.0%), and need for special therapy (4.3%; data not shown). Prevalence of SHCNs increased markedly with worsening parent-rated health, from 10.5% of children whose health was rated “excellent” to 62.5% of those whose health was rated “poor.” The fact that CSHCN by definition excludes acute health conditions (capturing only those lasting \( \geq 12 \) months) may explain why one third of children in poor parent-rated health did not have any SHCN. Asthma was associated with a sixfold greater risk of SHCN.

The vast majority (88%) of children remained enrolled in NJFC at the time of the survey, with \( \sim 7\% \) disenrolled but with other health insurance and 6% disenrolled and uninsured. CSHCN were more likely than non-CSHCN to remain enrolled in SCHIP (91.5% and 86.5%, respectively) and less likely to become uninsured (2% of CSHCN vs 7% of non-CSHCN). Approximately the same share of each group found other health insurance (6.3% and 6.6%).

Table 2 presents odds ratios (ORs) and 95% confidence intervals (CIs) from multinomial logit models of children’s final enrollment status (either disenrolled but still insured, or disenrolled and uninsured when each is compared with remaining enrolled in NJFC), controlling for age, gender, race/ethnicity, and NJFC plan level. CSHCN had only one fourth the odds of being disenrolled and uninsured compared with non-CSHCN. Hispanic children were \( \sim 3 \) times as likely as non-Hispanic white children to become uninsured. Children rated in fair or poor parent-rated health had 4 times the odds of being disenrolled and uninsured (data not shown), suggesting that those with acute health problems not captured by the SHCN screen had very different patterns than those with chronic conditions.

**DISCUSSION**

Our analysis of survey data of children from New Jersey’s SCHIP suggests that children with chronic health conditions seem to remain in SCHIP longer than healthy children. The data provide evidence of differential retention, with a considerably lower chance of becoming disenrolled and uninsured among CSHCN than among their healthier counterparts. CSHCN had only one fourth the chances of being disenrolled from NJFC and uninsured as children without such needs, even when demographic characteristics and NJFC plan level were taken into account. However, CSHCN had similar odds of disenrolling from the NJFC program and obtaining other insurance as children without special needs. Consequently, CSHCN were more likely than non-CSHCN to

<table>
<thead>
<tr>
<th>Variable</th>
<th>Enrolled</th>
<th>Disenrolled But With Other Health Insurance</th>
<th>Disenrolled and Uninsured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any SHCN, OR (95% CI)</td>
<td>0.82 (0.38–1.77)</td>
<td>0.25 (0.07–0.90)</td>
<td></td>
</tr>
<tr>
<td>Age groups, OR (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0- to 5-y-olds</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>6- to 12-y-olds</td>
<td>2.09 (0.76–5.81)</td>
<td>1.66 (0.57–4.80)</td>
<td></td>
</tr>
<tr>
<td>13- to 18-y-olds</td>
<td>1.41 (0.48–4.12)</td>
<td>1.27 (0.42–3.80)</td>
<td></td>
</tr>
<tr>
<td>Gender, OR (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girl</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Boy</td>
<td>1.33 (0.69–2.56)</td>
<td>1.58 (0.77–3.22)</td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity, OR (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>1.26 (0.54–2.85)</td>
<td>1.62 (0.53–4.95)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.64 (0.30–1.37)</td>
<td>3.10 (1.40–6.91)</td>
<td></td>
</tr>
<tr>
<td>Other race</td>
<td>0.40 (0.09–1.73)</td>
<td>0.93 (0.18–4.80)</td>
<td></td>
</tr>
<tr>
<td>NJFC plan level, OR (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plan A (( \leq 133% ) FPL)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Plan B/C (133%–200% FPL)</td>
<td>1.69 (0.78–3.67)</td>
<td>0.88 (0.40–1.95)</td>
<td></td>
</tr>
<tr>
<td>Plan D (201%–350% FPL)</td>
<td>2.63 (1.11–6.22)</td>
<td>1.16 (0.45–2.99)</td>
<td></td>
</tr>
<tr>
<td>–2 log-likelihood</td>
<td>283.71</td>
<td>283.71</td>
<td></td>
</tr>
</tbody>
</table>

Children still enrolled in NJFC are the reference group for all of the models. Data were weighted to the universe of all of the children enrolled in the NJFC program as of May 2002.
remain covered either by SCHIP or other health insurance at the time of the survey (98% of CSHCN and 93% of non-CSHCN). To our knowledge, ours is the first study to analyze patterns of subsequent insurance status among CSHCN who disenroll from SCHIP.

A previous analysis of administrative data from New Jersey’s SCHIP initiative showed that reasons for disenrollment included failure to renew at the time of re-determination of eligibility, nonpayment of premium, and moving away. 

Disenrollment does not seem to be because of dissatisfaction with the program, because most respondents were satisfied with the NJFC program, with >82% rating the program “excellent,” “very good,” or “good,” regardless of their final enrollment status.

The prevalence of CSHCN in NJFC, ~1 in 5 children, was higher than the 13% prevalence rates in the general child population at both state and national levels based on the 2001 National Survey of Children With Special Health Care Needs. 

However, the prevalence of CSHCN in NJFC was within the range found in studies of SCHIP programs in other states that used the same approach to measuring CSHCN. 

A variety of approaches have been used to identify CSHCN, causing some confusion in the use of that terminology and acronym, as well as associated variation in prevalence estimates. The current study and others compared above use a survey approach, asking the child’s parents ≥1 question from the CSHCN screener or the Questionnaire for Identifying Children With Chronic Conditions, both noncategorical approaches to classifying SHCNs. For example, an Urban Institute study compared several states, each of which used different variants of the screener approach on SCHIP applications or relied on health care providers to identify CSHCN. They report CSHCN prevalence rates ranging from 1% to 8% of enrolled children but also conclude that “states are dissatisfied with the ability of their systems to identify CSHCN,” particularly with a single-question screener approach on the application.

Other studies use International Classification of Diseases, Ninth Revision (ICD-9), codes on administrative claims or encounter data to identify minor, moderate, and major chronic conditions, which they classify as SHCN, often yielding lower estimates of CSHCN prevalence than those using the CSHCN screener. For instance, a study that measured disabilities and chronic health conditions using clinical risk groups based on the ICD-9, codes from claims and encounter records found that <10% of children enrolled in Medicaid or SCHIP had SHCN.

Consistent with patterns observed in the general population and in SCHIPs in New York, Kansas, and other states, we found that older children and boys were more likely to have a SHCN but that there was no significant variation in prevalence by family income. SHCNs were more common among children with asthma and with poor parent-rated health, helping to validate the CSHCN screener as an indicator of chronic health problems. The NJFC program has different levels of coverage for different income levels, and the coverage at higher income levels does not provide the same number and range of benefits as the Medicaid expansion plan. We found that children in the highest plan level, which covers families from 201% to 350% of the FPL, were more likely to obtain other insurance than those from lower-income families on disenrolling from NJFC, perhaps because families in higher-income ranges are more likely to transition to employer-based health insurance.

This study has several notable advantages for the analysis of differential health retention in SCHIP. First is the validity of the indicators of SHCNs, which was highly correlated with other health measures. The CSHCN screener captures a wider range of chronic conditions than surveys that ask only about specific health conditions. Second, the availability of 2 sources to verify enrollment status at the time of the survey allowed the survey responses to be checked against NJFC administrative records to determine whether the index child remained enrolled. Third, using the combined survey and administrative data allowed us to determine whether those who disenrolled found other health insurance and to assess continued eligibility for the NJFC program among those disenrolled based on family income and age composition.

A limitation of the study is its cross-sectional design, which means that we cannot determine SHCN status at the time the child enrolled. However, the chronic nature of the health conditions screened for suggests that those needs are relatively unlikely to change during the year between enrollment and the survey. Second, this study was limited to 1 state with comparatively generous SCHIP eligibility. Finally, the survey response rate was 52%, typical for a telephone survey of low-income families, but raising the possibility that the sample was not representative of all NJFC enrollees. The lower response rate among families that had disenrolled by the time of the survey suggests that we may have underestimated overall disenrollment rates, but there is no evidence to suggest that our estimates of differential retention of CSHCN would have been biased.

Implications for Policy

The policy implications of these results can be viewed from several perspectives, including the families of CSHCN, children’s health advocates, clinicians, and SCHIP administrators. First, families who have children with special needs seem to understand the importance of coverage for their children. It is reassuring that most CSHCN are insured, either publicly or privately; however, more research is needed to understand the mechanisms underlying the higher retention of CSHCN in SCHIP, including satisfaction with the program, use of health care services, and experiences with the renewal.
process. Second, child health advocates and clinicians will be pleased that children with special needs will have much lower odds of being disenrolled from SCHIP and uninsured, which increases the likelihood of coordinated, comprehensive care for these children. Finally, SCHIP administrators will also be heartened to know that this vulnerable population remains insured, because the goal of SCHIP is to provide coverage for children who would not otherwise have health insurance.

However, because CSHCN, by definition, use more health services than children without special needs, the higher retention of CSHCN in SCHIP implies greater costs associated with serving this unhealthier population of children. Although CSHCN constitute only a small proportion of enrolled children, they generate a disproportionate amount of expenses incurred. One study of 2 states’ SCHIP plans suggests that CHSCN made up ~15% of the total enrollment but accounted for ~60% of program expenditures.

As part of the federal reauthorization of SCHIP in 2007, Congress has the opportunity to reward states for maintaining high retention levels of CSHCN by adjusting the allotment of funds to account for states with higher proportions of children with special needs. States use a variety of strategies for financing care for CSHCN, including adjusting risk for demographic factors and health status, using carve-outs for specific health conditions, and reinsurance for children whose annual expenses exceed some threshold. A recent study of these strategies by the National Center on Financing for CSHCN concluded that reinsurance combined with risk adjustment for health status seemed to be the best strategy for aligning costs and payments for CSHCN. These strategies acknowledge the higher cost of caring for children who remain covered over time and reward plans for assuring that they are well served. Unfortunately, the anticipated federal shortfall for SCHIP funding of $10 billion to $12 billion for 2008–2012 raises substantial concerns about the program’s ability to meet these needs.

**Directions for Future Research**

We found that the majority of CSHCN enrolled in SCHIP retain that coverage; it would be useful to extend that research to investigate how well the program satisfies continuity and adequacy of coverage, the other 2 components of the health insurance core outcome developed by the Maternal and Child Health Bureau. Reaching the health insurance core outcome is crucial for children with special needs, because they use more health services than other children. Determining the health care use, unmet needs, and financial burden on families with CSHCN enrolled SCHIP will provide a more detailed evaluation of how well the program is serving this vulnerable population.

Future research should also investigate the prevalence and costs of the 5 types of health care needs encompassed in the SHCN screener (medication use, greater than average health service use, activity limitations, emotional problems, and need for special therapy). More detailed estimates of the components of costs associated with providing care for CSHCN within SCHIP would help ensure the long-term financial sustainability of SCHIP.

**ACKNOWLEDGMENTS**

This study was conducted in cooperation with the New Jersey Department of Human Services, which provided study data and invaluable advice. Partial funding for this project was provided by the US Department of Health and Human Services, Health Services and Research Administration, through a grant to the New Jersey Department of Human Services and Rutgers Center for State Health Policy.

We thank Susan Brownlee for assistance with the survey design and implementation and Michelle Walsky and Heidi Smith for advice and assistance.

**REFERENCES**


PEDIATRICS Volume 120, Number 5, November 2007 e1223


APPENDIX.
Sampling plan: 2003 NJFCS. NA indicates not applicable.

a Parental eligibility includes plans A, B, and C only.

b Enrollment status is as of December 31, 2002.
Impact of Size at Birth on the Microvasculature: The Avon Longitudinal Study of Parents and Children

Robyn J. Tapp, PhD\textsuperscript{a,b}, Cathy Williams, PhD\textsuperscript{c}, Nicholas Witt, MSc\textsuperscript{a}, Nish Chaturvedi, MD\textsuperscript{a}, Richard Evans, BA\textsuperscript{a}, Simon A. McG Thom, MD\textsuperscript{a}, Alun D. Hughes, MD\textsuperscript{a}, Andrew Ness, PhD\textsuperscript{c}

\textsuperscript{a}International Centre for Circulatory Health, National Heart and Lung Institute, St Mary’s Hospital and Imperial College London, United Kingdom; \textsuperscript{b}Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia; \textsuperscript{c}Department of Social Medicine, University of Bristol, United Kingdom

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

BACKGROUND. The impact of early life factors on the microvasculature is relatively unknown.

OBJECTIVES. We hypothesized that small birth size may be associated with structural variations in the retinal vasculature in children.

METHODS. The Avon Longitudinal Study of Parents and Children followed a cohort of children born in 1991–1992 from birth. The current study included the first 263 children who were systematically screened in the year-12 follow-up. Complete data were available for 166 children with a gestation of \( \geq 37 \) weeks. Retinal circulatory measures were evaluated, including retinal microvascular tortuosity and bifurcation optimality deviance, an indicator of abnormal endothelial function.

RESULTS. Optimality deviance and retinal tortuosity were higher among those with lower birth weight. Linear regression modeling was used to assess the association of retinal microvascular measures with birth weight. The standardized \( \beta \) coefficient between optimality deviance and birth weight was \(-.182\) adjusted for gender and age in weeks; additional adjustment for systolic blood pressure and heart rate had little impact on the \( \beta \) coefficient. A similar association was observed for retinal tortuosity.

CONCLUSION. The findings of this study suggest that early life factors may have an important impact on retinal vascular structure, possibly through an adverse effect on endothelial function.
The role of the microcirculation in the pathogenesis of cardiovascular disease (CVD) is being increasingly recognized. Retinal microvascular abnormalities have been associated with an increased risk of stroke, ischemic heart disease, and diabetes, independent of other well-established risk factors in adults. These are important observations that could further our understanding of mechanisms linking the long-observed association among low birth weight and CVD, hypertension, and diabetes in adults. Determining whether these associations represent cause or effect has been difficult in adult populations where these conditions are common. A population of children where the frequency of hypertension and other metabolic disorders is low provides the ideal setting to determine whether retinal microvascular abnormalities are associated with early life factors. We hypothesized that small birth size may be associated with structural variations in the retinal vasculature in children.

METHODS
The population, methods, and response rate for the Avon Longitudinal Study of Parents and Children (ALSPAC) are described in detail elsewhere. In brief, ALSPAC researchers enrolled pregnant women living in 3 health districts centered in Bristol, England, who had an expected date of delivery between the start of April 1991 and the end of December 1992. Of these, 11,211 had a white singleton live-born child. The current study includes the first 263 children systematically screened in the ALSPAC year-12 follow-up. The children were aged 12 years at this follow-up. Complete data were available for 166 children with a gestation of ≥37 weeks. There were no significant differences in the ages or genders (P = .269) of the youth with complete and incomplete data.

The date of the last menstrual period as reported by the mother at enrollment and the date of delivery were used to estimate gestation. Term birth was defined as birth on or after 37 completed weeks of gestation. Infant gender and birth weight were recorded in the delivery room and abstracted from obstetric records and/or birth notifications. Measures of childhood weight, height, heart rate, and blood pressure (BP) used in the present study were taken at age 9 years. Height was measured with shoes and socks removed using a Harpenden stadiometer (Holtain Ltd, Crymych, Pembs, United Kingdom) to the nearest 0.1 cm, and weight was measured by using a Tanita TBF 305 body-fat analyzer and weighing scales (Tanita UK Ltd, Yewley, Middlesex, United Kingdom). BMI was calculated as weight (kilograms)/height (meters squared). Systolic and diastolic BP were recorded on the right arm while the subject was seated using a Dinamap 9301 vital signs monitor (Critikon, Tampa, FL).

We took 45° digital retinal images of the macular center of each retina using a Topcon nonmydriatic retinal camera (Topcon TRC-NW6s, Topcon Technologies, Paramus, NJ) fitted with a Nikon D1X (Nikon, Tokyo, Japan). Images were graded by 2 observers who were blinded to subject data using a semiautomated system that captures a wide range of retinal geometric parameters, and reproducibility of this technique has been reported previously. Measured parameters included the (1) arteriolar diameters, (2) arteriolar bifurcation angles, (3) length/diameter ratios of arteriolar segments and arteriolar/venular diameter ratios (these parameters provide measures of arteriolar narrowing that are relatively unaffected by differences in optical refraction), (4) arteriolar tortuosity (estimated as the actual length of the vessel divided by the straight line distance between bifurcations minus 1), and (5) arteriolar optimality ratio and optimality deviance. Optimality ratio is the ratio of sum of “daughter” arteriolar diameters divided by the “parent” arteriolar diameter corrected for asymmetry. For a theoretically optimal bifurcation, the optimality ratio should be 0.79, and the optimality deviance was calculated as the absolute value of the optimality ratio minus 0.79.

Ethical approval of the study was obtained from the ALSPAC Law and Ethics Committee. Written informed consent for the study was obtained.

Statistical Methods
The data analysis was performed with SPSS 14.0 for Windows (SPSS Inc, Chicago, IL). Descriptive information for each of the variables was derived, and the distribution was assessed. Baseline data are presented as mean ± SE or percentages. Linear regression was used to assess the association of birth weight with measures of the retinal microcirculation. Standardized β coefficients were used, because they allow for direct comparison of the strength of associations between risk factors and disease.

RESULTS
The characteristics of the population are shown in Table 1. Optimality deviance and retinal tortuosity were higher among those with lower birth weight (birth weight groups: <3.2, 3.2–3.6, and >3.6 kg).

Linear regression modeling was used to assess the association of retinal vascular measures with birth weight (Table 2). Linear regression identified lower birth weight as an independent factor associated with increased optimality deviance after adjustment for age (in weeks), gender, and systolic BP. The standardized β coefficient between optimality deviance and birth weight was −.208 (P = .007); adjustment for gender, age in weeks, systolic BP, and heart rate had little impact on the β coefficient (β = −.184; P = .021). Further adjustment for BMI at age 9 years had little impact on...
TABLE 1 Characteristics of the Population at Age 9 Years According to Birth Weight Group

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Birth Weight &lt; 3.2 kg</th>
<th>Birth Weight 3.2–3.6 kg</th>
<th>Birth Weight &gt; 3.6 kg</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>50</td>
<td>69</td>
<td>47</td>
<td>.004</td>
</tr>
<tr>
<td>Female, %</td>
<td>35</td>
<td>46</td>
<td>19</td>
<td>.084</td>
</tr>
<tr>
<td>BMI, mean ± SE, kg/m²</td>
<td>16.8 ± 0.4</td>
<td>17.4 ± 0.3</td>
<td>18.1 ± 0.2</td>
<td>.020</td>
</tr>
<tr>
<td>Height, mean ± SE, cm</td>
<td>137.8 ± 0.9</td>
<td>140.4 ± 0.8</td>
<td>141.2 ± 0.9</td>
<td>.021</td>
</tr>
<tr>
<td>Weight, mean ± SE, kg</td>
<td>32.1 ± 0.9</td>
<td>34.5 ± 0.9</td>
<td>36.3 ± 1.2</td>
<td>.581</td>
</tr>
<tr>
<td>Systolic BP, mean ± SE, mm Hg</td>
<td>103.6 ± 1.3</td>
<td>105.5 ± 1.2</td>
<td>104.6 ± 1.6</td>
<td>.984</td>
</tr>
<tr>
<td>Diastolic BP, mean ± SE, mm Hg</td>
<td>58.9 ± 0.9</td>
<td>58.7 ± 0.7</td>
<td>58.7 ± 0.9</td>
<td>.552</td>
</tr>
<tr>
<td>Heart rate, mean ± SE, bpm</td>
<td>80.2 ± 1.4</td>
<td>78.2 ± 1.2</td>
<td>79.6 ± 1.4</td>
<td>.132</td>
</tr>
<tr>
<td>Arteriolar optimality deviance, mean ± SE</td>
<td>0.094 ± 0.01</td>
<td>0.076 ± 0.007</td>
<td>0.070 ± 0.007</td>
<td>.375</td>
</tr>
<tr>
<td>Arteriolar bifurcation angle, mean ± SE</td>
<td>8.0 ± 1.9</td>
<td>7.96 ± 1.7</td>
<td>7.63 ± 1.2</td>
<td>.031</td>
</tr>
<tr>
<td>Arteriolar simple tortuosity, mean ± SE</td>
<td>0.031 ± 0.004</td>
<td>0.025 ± 0.003</td>
<td>0.018 ± 0.002</td>
<td>.470</td>
</tr>
<tr>
<td>Arteriolar LDR, mean ± SE</td>
<td>13.1 ± 0.5</td>
<td>13.5 ± 0.4</td>
<td>12.7 ± 0.5</td>
<td>.873</td>
</tr>
<tr>
<td>Arteriolar diameter, mean ± SE, pixels</td>
<td>15.7 ± 1.5</td>
<td>15.7 ± 1.4</td>
<td>15.8 ± 1.5</td>
<td>.310</td>
</tr>
<tr>
<td>Venular diameter, mean ± SE, pixels</td>
<td>20.1 ± 3.7</td>
<td>19.5 ± 3.3</td>
<td>19.1 ± 3.2</td>
<td>.584</td>
</tr>
<tr>
<td>AVR, mean ± SE</td>
<td>0.90 ± 0.02</td>
<td>0.91 ± 0.02</td>
<td>0.91 ± 0.01</td>
<td>.654</td>
</tr>
</tbody>
</table>

LDR indicates length/diameter ratio; AVR, arteriolar/venular diameter ratio; bpm, beats per minute.

TABLE 2 Standardized β Coefficients and P Values for Measures of the Retinal Vasculature With Population Characteristics at Age 9 Years

<table>
<thead>
<tr>
<th>Variable</th>
<th>Arteriolar Optimality Deviance</th>
<th>Arteriolar Simple Tortuosity</th>
<th>Arteriolar Bifurcation Angle</th>
<th>Arteriolar LDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>P</td>
<td>β</td>
<td>P</td>
<td>β</td>
</tr>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, kg</td>
<td>−182</td>
<td>0.023</td>
<td>−198</td>
<td>0.013</td>
</tr>
<tr>
<td>Gender</td>
<td>0.064</td>
<td>0.424</td>
<td>0.050</td>
<td>0.524</td>
</tr>
<tr>
<td>Age, wk</td>
<td>−0.001</td>
<td>0.989</td>
<td>0.087</td>
<td>0.258</td>
</tr>
<tr>
<td>Model 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, kg</td>
<td>−184</td>
<td>0.021</td>
<td>−193</td>
<td>0.014</td>
</tr>
<tr>
<td>Gender</td>
<td>0.050</td>
<td>0.539</td>
<td>0.087</td>
<td>0.273</td>
</tr>
<tr>
<td>Age, wk</td>
<td>−0.011</td>
<td>0.892</td>
<td>0.068</td>
<td>0.374</td>
</tr>
<tr>
<td>Systolic BP, mm Hg</td>
<td>0.089</td>
<td>0.273</td>
<td>0.056</td>
<td>0.480</td>
</tr>
<tr>
<td>Heart rate, bpm</td>
<td>0.040</td>
<td>0.629</td>
<td>−0.193</td>
<td>0.017</td>
</tr>
<tr>
<td>Model 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, kg</td>
<td>−190</td>
<td>0.021</td>
<td>−179</td>
<td>0.026</td>
</tr>
<tr>
<td>Gender</td>
<td>0.048</td>
<td>0.554</td>
<td>0.091</td>
<td>0.255</td>
</tr>
<tr>
<td>Age, wk</td>
<td>−0.018</td>
<td>0.823</td>
<td>0.085</td>
<td>0.288</td>
</tr>
<tr>
<td>Systolic BP, mm Hg</td>
<td>0.074</td>
<td>0.418</td>
<td>0.088</td>
<td>0.325</td>
</tr>
<tr>
<td>Heart rate, bpm</td>
<td>0.038</td>
<td>0.639</td>
<td>−0.190</td>
<td>0.019</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>0.033</td>
<td>0.728</td>
<td>−0.072</td>
<td>0.437</td>
</tr>
</tbody>
</table>

bpm indicates beats per minute.

the association with birth weight. A similar association was observed for retinal tortuosity (Table 2).

DISCUSSION
To our knowledge, this is one of the first studies to assess the relationship between birth weight and the retinal microvasculature in children. We have shown that low birth weight is strongly associated with increased optimality deviance and retinal tortuosity, findings that fit well with the known association among low birth weight, stroke, and CVD in adults and we have provided some insights into possible mechanisms underlying these associations.

The present study suggests that early life factors may have an important impact on microvascular structure and function in children. Optimality deviance is an indicator of endothelial dysfunction, and this may imply that a primary disorder of the endothelium in early life is the mechanistic link between birth weight and CVD. This is consistent with previous findings of impaired endothelial function in large conduit arteries in children and young adults of low birth weight and decreased capillary recruitment and reduced vasodilator responses to acetylcholine in the skin of children of low birth weight. Although this question has not been examined previously in the retinal microvasculature of children, Chapman et al observed that adult men of low birth weight had narrower bifurcation angles compared with normal birth weight control subjects, indicative of retinal arteriolar rarefaction. Similarly, in a study by Kistner et al of women born preterm, the researchers observed a higher length index for arterioles and a decreased num-
ber of vascular branching points compared with women born at term. Both of these studies concluded that the observed architectural changes in the retinal microvasculature might be related to impaired fetal growth. The present study was small and underpowered to detect some associations. Further investigation is warranted in a well-documented large prospective study of children at risk of future vascular disease.

CONCLUSION
The findings of this study suggest that early life factors may have an important impact on vascular structure, possibly through an adverse effect on endothelial function.

ACKNOWLEDGMENTS
The United Kingdom Medical Research Council, the Wellcome Trust, and the University of Bristol provide core support for the ALSPAC. Dr Tapp is supported by a Sidney Sax Fellowship from the National Health and Medical Research Council of Australia (grant 334173). Dr Witt is supported by a fellowship from the Foundation for Circulatory Health.

We are extremely grateful to all of the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses.

REFERENCES
OBJECTIVES. Among a cohort of long-term neuroblastoma survivors, our aims were to (1) assess the association between treatment intensity and parent-reported hearing loss in the child, (2) evaluate the strength of the association between hearing loss and parent-reported academic and psychosocial difficulties in the child, and (3) examine the association between parent-reported academic and psychosocial difficulties in the child and the child’s self-reported quality of life.

PATIENTS AND METHODS. Through a mailed survey that included the Pediatric Quality of Life Inventory 4.0 and an outcomes questionnaire for parents, we evaluated 137 children (aged 8–17 years) who were previously enrolled in 1 of 2 Children’s Cancer Group neuroblastoma clinical studies.

RESULTS. Childhood survivors of neuroblastoma who had prevalent hearing loss, as reported by their parents, had at least twice the risk of an identified problem with reading skills, math skills, and/or attention and a similarly higher risk of a general learning disability and/or special educational needs than did neuroblastoma survivors without hearing loss. Consistent with this finding, hearing loss was associated with a 10-point-lower mean score in the school-functioning scale of the Pediatric Quality of Life Inventory 4.0. We also observed a clear pattern of poorer self-reported quality-of-life scores among children with parent-reported academic and psychosocial problems compared with those without such problems, particularly with school functioning, even after controlling for reported hearing loss.

CONCLUSIONS. We found evidence that long-term neuroblastoma survivors, especially those with hearing loss, are at elevated risk for academic learning problems and psychosocial difficulties. We also found strong concordance between parent-reported learning problems in the child and indications of distress in the child’s self-reported quality of life.
Neuroblastoma is an embryonal malignancy of the sympathetic nervous system that predominantly occurs in infants and very young children; the median age of diagnosis is 19 months. The adrenal gland is the most common tumor location of origin, but neuroblastoma tumors may arise almost anywhere in the body, including the brain or spinal cord. Prognosis varies widely depending on risk stratification at diagnosis, with stage; age; myc myelocytomatosis viral–related oncogene, neuroblastoma derived (MYCN) status; DNA index; and histopathology as important predictors of survival probability.

There is evidence that survivors of neuroblastoma may experience a number of long-term adverse effects from the tumor and its treatment. Hearing loss is one important potential treatment-related late effect, caused usually from exposure to platinum-containing chemotherapies, particularly cisplatin and myeloablative doses of carboplatin. High cumulative doses of platinum chemotherapy are included in multimodal treatment protocols for intermediate- and high-risk neuroblastoma. Platinum-related ototoxicity results from destruction of cochlear sensory hair cells. Because these specialized sensory cells are arranged tonotopically (in order of pitch), the initial hearing loss associated with platinum-based chemotherapy is generally in the high-frequency ranges (>2000 Hz) and results from sensory hair cell destruction that begins at the base of the cochlea and continues toward the cochlear apex, where the speech frequencies (500–2000 Hz) are impacted as cumulative doses of platinum chemotherapy increase. Even minimal-to-mild hearing loss in the high-frequency ranges can significantly impact language development, verbal abilities, and reasoning skills in young children. This is of particular concern in children with neuroblastoma who receive platinum-based chemotherapy, because they are often still in the process of acquiring language at the time of their treatment. Many of these children will experience therapy-related hearing loss initially affecting the high-frequency ranges. The impact of this high-frequency hearing loss on language development in young children is important and often underappreciated.

The high-frequency speech phonemes are critical to speech intelligibility and the identification of verb tense and plural forms. The English language relies heavily on fricative consonants (sounds such as “s,” “f,” and “th,” which are produced by forcing air through a narrowing of the vocal tract) to convey the meaning of words. Because fricative sounds are primarily high frequency in nature, even mild-to-minimal losses in the high-frequency range have been associated with problems with language acquisition and speech discrimination, which are associated with increased academic and social-emotional difficulties in young children. Given the important age-dependent developmental tasks of speech and language acquisition and the recognition that these tasks mirror the usual age at neuroblastoma treatment, we were concerned about the potential effects on psychosocial and academic functioning among childhood survivors of neuroblastoma. The purpose of this report is to disseminate findings from a cross-sectional follow-up analysis of quality of life (QoL) among 137 survivors of neuroblastoma who were enrolled on 1 of 2 national clinical studies through the Children’s Cancer Group (now part of the Children’s Oncology Group [COG]). The aims of the analysis were threefold: (1) to describe the association between treatment intensity and parent-reported hearing loss in the child; (2) to evaluate the strength of the association between hearing loss and parent-reported academic and psychosocial difficulties in the child; and (3) to examine the association between parent-reported academic and psychosocial difficulties in the child and the child’s self-reported QoL.

**METHODS**

**Patients and Recruitment**

Potential subjects included children enrolled previously in Children’s Cancer Group studies CCG3881 or CCG3891. CCG3881 was a nonrandomized study of children diagnosed before 19 years of age with low- or intermediate-risk neuroblastoma between 1989 and 1995. CCG3891 was a randomized clinical trial comparing intensive chemotherapy (standard of care) versus the addition of autologous bone marrow transplantation for children diagnosed with high-risk neuroblastoma between 1991 and 1996. Children treated on CCG3881 were classified as low risk if they had stage I, II, or IVS disease; they received surgery and supportive care, with chemotherapy (cyclophosphamide) given only for symptomatic stage IVS. Fewer than 15% of patients received combination chemotherapy as a salvage for recurrence. Patients were classified as intermediate risk if they had stage III disease at <1 year of age, were ≥1 year of age with stage III and favorable biology (MYCN nonamplified, ferritin <143 ng/mL, favorable Shimada), or were <1 year of age with stage IV disease and no MYCN amplification. These patients received ~9 months of combination therapy with cisplatin (total dose: 510 mg/m²), cyclophosphamide, doxorubicin, and etoposide. High-risk patients were treated on CCG3891 with either intensive induction chemotherapy for 5 months with cisplatin (300 mg/m²), cyclophosphamide, etoposide, and doxorubicin, followed by either 3 additional chemotherapy cycles with cisplatin (480 mg/m²), ifosfamide, etoposide, and doxorubicin or with myeloablative chemotherapy with carboplatin (1000 mg/m²), etoposide, and melphalan, with total body irradiation and autologous bone marrow transplant.

We attempted to secure institutional participation for this follow-up study, including human subjects (institu-
tional review board [IRB]) approval, from each of the 99 treating hospitals with ≥1 previously enrolled child on CCG3881 and or CCG3891 who were believed to be still alive when the current study began. We were successful in obtaining local IRB approval to open the study in 63 treating institutions. In most cases, local IRB procedures required the treating institution to obtain consent from the family before allowing us to contact the parents to offer study participation. The COG statistical office sent each treating institution a list of their potentially eligible patients for this purpose (the Children’s Cancer Group was merged into COG during the early course of this study). Of the 843 children identified through COG who were alive at last contact and potentially eligible, we received permission from the treating institutions to contact the parents of 253 children. Of these, 10 did not meet study eligibility (5 were older than 18 years, 4 were from non–English-speaking households, and in 1 we could not clearly determine treatment status). Of the 243 eligible, 44 (18.1%) were lost to follow-up and never contacted, 8 (3.3%) actively refused, 51 (21%) did not respond to requests to participate or did not return study materials, 3 (1.2%) had parental but not child participation, and 137 children with 1 or both of their parents participated in the study after providing informed consent (68.8% of the 199 contacted; 56.4% of the 243 presumed eligible with permission to attempt contact). In addition to being previously enrolled on either CCG3881 or CCG3891, eligibility criteria required children to be in cancer remission, ≤18 years of age when completing the QoL questionnaire, and to have participation from 1 or both parents with completion of a parent questionnaire (provided only in English).

Data Collection
Participating children completed the Pediatric Quality of Life Inventory 4.0 (PedsQL) Generic Core scales (www.pedsqql.org), which is a global QoL instrument with established reliability, validity, and normative values in both healthy and clinical (normal) populations. The PedsQL has been used for studies in childhood cancer survivors and in school populations. The generic scales of the 23-item PedsQL measure functioning in physical, emotional, social, and school domains, with 2 summary scores considered here: psychosocial and total.

Participating parents completed an instrument devised by our clinical study team that asked simple questions on late effects after neuroblastoma treatment, including ones related to select medical problems, academic problems, and behavioral or mental health problems. The “Neuroblastoma Medical Follow-up” section of the questionnaire had directions that read, “Have you been told by a doctor or other health care professional that your child, who had neuroblastoma, currently has (or had in the past) any of the following.” Response categories were “no”; “yes, in the past, but not now”; and “yes, and it continues.” A comment field was also provided for parental response. Responses to “yes, in the past, but not now” were too few to evaluate separately, so they were combined with the other “yes” category for this analysis. We did not ask questions on severity or type of hearing loss or similar information for any of the other medical or psychosocial outcomes. No validation was made of the information provided by the parents. There were 83 parent pairs and 54 lone parents (52 mothers and 2 fathers) who completed the parent questionnaire. For parent dyads, each parent was instructed to complete the questionnaire independently; we also asked that the child complete the PedsQL instrument independent of his or her parents, although no effort to confirm compliance with these instructions was undertaken. A high degree of concordance between mother-father dyads was evident for each of the questions we address in this report, ranging from 94% for problems with writing skills and for behavioral concerns, to 91% for several of the measures, including problems with reading skills. Concordance was 93% for the hearing loss question. To maintain consistency of response, for this analysis only the mothers’ responses, except for the 2 single fathers, were considered. Basic data on diagnostic characteristics and treatment protocol were received from the treating institution that accompanied the information data sheet that they forwarded to us with permission to contact the family.

Statistical Procedures
To evaluate whether and to what extent differences were apparent between the 137 study participants and the 106 presumed eligible nonparticipants, descriptive statistics were calculated for demographic and treatment variables and compared with 2 sample t tests. Fisher’s exact test was used for categorical outcomes when cell sizes were <5. As described above for risk categorization, treatment intensity was categorized as low (CCG3881), intermediate (CCG3881), or high (CCG3891). The frequencies and percentages of parent-reported school problems and behavioral symptoms were also calculated to describe educational and psychosocial problems.

Raw values from responses to the Likert scale questions of the PedsQL were linearly transformed and for each scale were summed and divided by the number of questions to produce scores from 0 to 100. Means and SDs were compared between the participating children’s responses and a normative sample of children used in validation of the instrument using 2-sample t tests. QoL scores from study subjects were compared using stratified analysis by treatment intensity (3 levels), hearing loss (yes or no), current age (dichotomized as 8–12 and 13–17 to correspond with the age-specific PedsQL instruments), and gender. No adjustments for multiple comparisons were made.
A path model was constructed to evaluate the correlations and covariances between variables of interest and to guide the modeling of the effects of treatment intensity and hearing loss on academic and psychosocial outcomes and demographic characteristics. Based on the results of the path analysis (data not shown), multivariable logistic regression models were constructed to evaluate the strength of the associations between hearing loss and parental report of educational and psychosocial problems controlling for gender and to evaluate the associations between the children’s self-reported QoL and parent-reported academic and psychosocial problems controlling for gender and hearing loss.

RESULTS

Subject Characteristics

No statistical differences were observed between participants and presumed eligible nonparticipants with regard to age at diagnosis, study protocol, MYCN amplification, Evans stage at diagnosis, treatment intensity, or gender. Participants, on average, were somewhat younger than nonparticipants and had fewer years since diagnosis (Table 1). Among the 137 study subjects, mean age at diagnosis was 1.4 years, mean age at interview was 12.1 years (range: 8–17 years), and 49.6% were female. Twenty-five subjects received high treatment intensity, 54 subjects received intermediate treatment intensity, and 58 subjects received low treatment intensity.

Frequency of Outcomes

Hearing loss was reported in 43 (31.4%) of the 137 children. No difference in the percentage with hearing loss was observed between boys and girls (Table 2).

Parents were asked whether their child had “identified learning problems” with reading skills, writing skills, and math skills and whether their child had a general learning disability or special educational needs in school. Responses identifying a problem ranged from 31.4% for math-skill problems to 22.6% for a general learning disability (Table 2). Eighty children (58.4%) were reported as having no identified learning problem, 12 (8.8%) with 1 learning problem only, and 45 (32.8%) had ≥2 problems. Parents were also asked about whether their child had problems with poor attention, behavioral concerns, depression, and anxiety. Problem responses ranged from 27.7% for poor attention to 18.2% for depression (Table 2). Eight-five children (62.0%) were reported with none of the psychosocial problems, 19 (13.9%) with 1 problem only, and 33 children (24.1%) with ≥2 psychosocial problems.

Treatment Intensity and Hearing Loss

Hearing loss was reported in 52% of those who received high-intensity treatment, 44% of those who received intermediate-intensity treatment, and 10% of those who received low-intensity treatment. Relative to the low treatment intensity group, the odds of a reported hearing loss was 8.3-fold higher (95% confidence interval [CI]: 2.9–23.7) for children in the intermediate-intensity group and 9.9-fold higher (95% CI: 3.1–32.4) for children in the high-intensity group after controlling for gender (Table 3).
Hearing Loss and Parent-Reported Academic and Psychosocial Problems

Hearing loss was generally associated with increased parental report of academic problems in the child. In particular, identified problems with reading skills, math skills, poor attention, general learning disability, and special educational needs in school had odds ratios (ORs) of ≥2 in those with hearing loss compared with those without, after controlling for gender. Substantive differences by hearing loss were not observed for problems with writing skills, behavioral concerns, anxiety, or depression (Table 3).

Child-Reported QoL and Parent-Reported Academic and Psychosocial Problems

Mean scores were not statistically or substantively different in any QoL domain or in the QoL summary scores between study subjects and population norms. No substantive or statistical differences were observed in mean QoL among study subjects by treatment intensity or age at survey.

Boys generally reported poorer QoL scores than girls in school and social functioning and in the psychosocial and total QoL summary scores. After controlling for gender and hearing loss, strong and consistent associations between mean child-reported QoL score and parent-reported academic and psychosocial problems were observed. Except for physical functioning, within each category of academic and psychosocial functioning, the mean QoL score among children with a parent-reported problem was substantially and significantly lower than in children without an identified problem (Table 4).

**DISCUSSION**

We found that childhood survivors of neuroblastoma who had prevalent hearing loss, as reported by their parents, had at least twice the risk of an identified problem with reading skills, math skills, and/or attention and a similarly higher risk of a general learning disability and/or special educational needs than did survivors of neuroblastoma without hearing loss. Although it is possible that the associations between parent-reported hearing loss and academic and psychosocial problems in children are influenced to some extent by shared method variance, consistent with these results is a clear pattern of poorer self-reported QoL scores among children with parent-reported academic and psychosocial problems, particularly with school functioning, compared with those without such problems, even after controlling for reported hearing loss. These findings, the first of their kind to our knowledge, are important for parents, teachers, and health care providers to understand the potential impact on the child's psychosocial well-being many years after cancer therapy. Hearing loss, an important adverse late effect of treatment with platinum-based drugs, may be particularly problematic in neuroblastoma survivors because of the very young age at which they are treated, thus potentially adversely affecting speech and language development and subsequent academic performance. Cisplatin is included in some treatment protocols for many other childhood cancers, including medulloblastoma, osteosarcoma, and germ cell tumors, so these results may be informative beyond that of the neuroblastoma setting.

Our results add to the small but emerging literature on long-term outcomes related to neuroblastoma. Laverdiere et al evaluated medical morbidity in 63 survivors of advanced-stage neuroblastoma with a median follow-up time of 7 years after diagnosis. One or more late-
complications was detected in 95% of survivors, including hearing loss (62%), primary hypothyroidism (24%), ovarian failure (41% of females), musculoskeletal problems (19%), and pulmonary problems (19%). Barr et al. conducted a QoL study using a health utility index that compared 26 survivors of neuroblastoma to 52 Wilms’ tumor survivors with both groups averaging ~5 years since the end of treatment. Although QoL scores did not differ significantly between the 2 disease groups, morbidity burden, as reported by parents, was substantially greater for the neuroblastoma group than the Wilms’ tumor group, with hearing loss largely responsible for the difference. As part of the Childhood Cancer Survivor Study, Ness et al. compared limitations in physical performance and daily activities from a survey of 11,481 long-term survivors of childhood cancer (children and young adults; 88% were ≥10 years postdiagnosis), including 802 survivors of neuroblastoma, with 3839 siblings of childhood cancer survivors. They found limitations in physical performance reported in 17% of long-term survivors of neuroblastoma, a risk that was 70% higher than expected from the sibling comparison group. Elevated risks for restricted personal care skills (relative risk [RR]: 3.8; 95% CI: 2.2–6.8), restricted routine activities of daily living (RR: 3.6; 95% CI: 2.5–5.4), and problems with health that prevented school or work attendance (RR: 5.1; 95% CI: 3.4–7.6) were also observed. Ness et al. noted that, given the diagnosis period of the cohort (1970–1986), few high-risk neuroblastoma patients would have survived the required 5 years after diagnosis to be eligible for their study, so this neuroblastoma series, like ours, largely represented survivors of low- and intermediate-risk disease.

Although strengths of this study include the relatively large sample size and the inclusion of child self-reported QoL, several limitations need to be considered when interpreting the findings. Foremost, we were unable to confirm or characterize the type or severity of hearing loss reported, and we did not have data on any hearing treatments received, such as assisted devices or cochlear implants. Nor did we have data on actual school performance or confirmation of the psychosocial measures reported on by parents. However, the high degree of concordance between parent dyads for these measures, ranging from 91% to 94% agreement, is encouraging in relation to the validity of the responses that parents provided. Also of importance, because of our limited success in accessing and contacting potentially eligible parents through the many institutions that enrolled patients on the 2 neuroblastoma protocols, our study sam-

<table>
<thead>
<tr>
<th>Mean Child-Reported QoL Scores by Parent-Reported Academic and Psychosocial Problems Among Participating Survivors of Neuroblastoma (N = 137)</th>
<th>Physical</th>
<th>Emotional</th>
<th>Social</th>
<th>School</th>
<th>Psychosocial</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent-Reported Academic or Psychosocial Problems</td>
<td>Mean</td>
<td>P</td>
<td>Mean</td>
<td>P</td>
<td>Mean</td>
<td>P</td>
</tr>
<tr>
<td>Reading-skill problem</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (N = 95)</td>
<td>88.0</td>
<td>.13</td>
<td>81.4</td>
<td>.01</td>
<td>88.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Yes (N = 42)</td>
<td>84.1</td>
<td>—</td>
<td>71.2</td>
<td>—</td>
<td>76.3</td>
<td>—</td>
</tr>
<tr>
<td>Writing-skill problem</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (N = 101)</td>
<td>88.4</td>
<td>.02</td>
<td>81.0</td>
<td>.01</td>
<td>87.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Yes (N = 36)</td>
<td>82.1</td>
<td>—</td>
<td>70.6</td>
<td>—</td>
<td>74.2</td>
<td>—</td>
</tr>
<tr>
<td>Math-skill problem</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (N = 94)</td>
<td>87.6</td>
<td>.27</td>
<td>80.9</td>
<td>.02</td>
<td>87.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Yes (N = 43)</td>
<td>84.8</td>
<td>—</td>
<td>72.6</td>
<td>—</td>
<td>77.2</td>
<td>—</td>
</tr>
<tr>
<td>General learning disability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (N = 106)</td>
<td>87.8</td>
<td>.14</td>
<td>81.1</td>
<td>.005</td>
<td>87.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Yes (N = 31)</td>
<td>83.5</td>
<td>—</td>
<td>69.5</td>
<td>—</td>
<td>73.5</td>
<td>—</td>
</tr>
<tr>
<td>Special education need</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (N = 98)</td>
<td>87.7</td>
<td>.22</td>
<td>80.4</td>
<td>.05</td>
<td>87.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Yes (N = 39)</td>
<td>84.4</td>
<td>—</td>
<td>73.0</td>
<td>—</td>
<td>76.0</td>
<td>—</td>
</tr>
<tr>
<td>Poor attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (N = 99)</td>
<td>88.3</td>
<td>.07</td>
<td>81.9</td>
<td>.003</td>
<td>88.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Yes (N = 38)</td>
<td>83.3</td>
<td>—</td>
<td>70.1</td>
<td>—</td>
<td>75.5</td>
<td>—</td>
</tr>
<tr>
<td>Behavioral concerns</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (N = 110)</td>
<td>87.4</td>
<td>.22</td>
<td>80.7</td>
<td>.003</td>
<td>86.0</td>
<td>.01</td>
</tr>
<tr>
<td>Yes (N = 27)</td>
<td>83.7</td>
<td>—</td>
<td>68.2</td>
<td>—</td>
<td>76.7</td>
<td>—</td>
</tr>
<tr>
<td>Depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (N = 112)</td>
<td>87.8</td>
<td>.05</td>
<td>81.8</td>
<td>&lt;.001</td>
<td>85.9</td>
<td>.01</td>
</tr>
<tr>
<td>Yes (N = 25)</td>
<td>81.7</td>
<td>—</td>
<td>62.5</td>
<td>—</td>
<td>76.8</td>
<td>—</td>
</tr>
<tr>
<td>Anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (N = 111)</td>
<td>87.8</td>
<td>.08</td>
<td>82.2</td>
<td>&lt;.001</td>
<td>86.1</td>
<td>.01</td>
</tr>
<tr>
<td>Yes (N = 26)</td>
<td>82.4</td>
<td>—</td>
<td>62.3</td>
<td>—</td>
<td>76.3</td>
<td>—</td>
</tr>
</tbody>
</table>

Means were adjusted for hearing loss and gender.
peple may not fully represent surviving patients. Among those who we were permitted to attempt to contact for study inclusion, no important differences were apparent between children who participated and those who did not. Unfortunately, we could not conduct a similar evaluation between the participating children and those children who we were not granted permission to contact from their treating institution.

CONCLUSIONS
We found evidence that long-term survivors of neuroblastoma, particularly those with hearing loss, are at elevated risk for academic learning problems and psychosocial difficulties. We also found strong concordance between parent-reported learning problems in the child and indications of distress as measured by the child’s self-reported QoL. Overall, however, the children’s self-reported QoL scores did not differ substantively from those of population norms, which is an encouraging finding that may reflect the resilience of these long-term childhood cancer survivors and their families.

ACKNOWLEDGMENTS
This work was funded by American Cancer Society grant RSG-01-194-01-PBP, Hope Street Kids grant HSK 052102, and National Cancer Institute grant CA13539. A complete listing of grant support for research conducted under study inclusion, no important differences were apparent between children who participated and those who did not. Unfortunately, we could not conduct a similar evaluation between the participating children and those children who we were not granted permission to contact from their treating institution.

REFERENCES


Size at Birth and Motor Activity During Stress in Children Aged 7 to 9 Years

Wolff Schlotz, PhD, Alexander Jones, PhD, Naomi M. M. Phillips, Keith M. Godfrey, PhD, David I. W. Phillips, PhD

Objective. Small size at birth is linked with metabolic and cardiovascular disease. There is increasing evidence that it is also linked with physiologic stress responses and abnormal behavior, in particular, symptoms of hyperactivity. Therefore, we investigated associations between size at birth and motor activity during psychosocial stress.

Methods. In 123 children aged 7 to 9 years, we examined the relations of birth weight, head circumference, length, and ponderal index at birth with motor activity on exposure to both stress and nonstress situations. Videos were recorded while the children performed a story and a math task in front of an audience (stress) and watched a movie (nonstress); motor activity was defined as lifting or tilting of a foot.

Results. Children who had a smaller head circumference at birth demonstrated greater motor activity during the stress test. There were marked gender differences in the results. In boys, lower birth weight, head circumference, and ponderal index were associated with greater motor activity during the stress test but not associated with motor activity during the nonstress situation. The findings remained significant when potential confounding variables were controlled for. There were no associations in girls.

Conclusions. The findings suggest long-term effects of an adverse fetal environment on the behavioral stress response in boys and parallel similar gender-specific effects on different stress response systems in humans and animals. The results could reflect permanent alterations of dopaminergic neurotransmission and have implications for the etiology of clinical hyperactivity.
Small size at birth is linked with metabolic and cardiovascular disease. There is increasing evidence that it is also linked with abnormal behavior, in particular, hyperactivity. In children, birth weight has been shown to be inversely related to symptoms of attention-deficit/hyperactivity disorder (ADHD), as well as various other behavioral problems.

However, only a few studies have tested associations of size at birth with biobehavioral stress responses in humans. In these studies, low birth weight has been associated with enhanced blood pressure and heart rate responses to psychological stressors in women but not men, whereas associations with enhanced hypothalamic-pituitary-adrenal axis responses to psychosocial stress have been observed in young men but not in girls. In a cohort of 18-year-old boys who underwent a standardized interview to diagnose suitability for military combat duty in Sweden, it has been shown that psychological stress susceptibility as assessed by the interviewer was inversely associated with birth weight and head circumference. These studies suggest that an adverse fetal environment may influence biobehavioral stress responses. However, no human study has investigated associations of size at birth with motor activity in response to a psychosocial stressor.

We observed motor activity during a psychosocial stress situation and during a nonstress situation in children aged 7 to 9 years whose mothers had been recruited in early pregnancy for a previous research study. On the background of associations of anthropometric birth outcome measures with hyperactivity and physiologic stress responses in children, we examined for associations between size at birth and motor activity during these situations.

**METHODS**

**Participants**
The mothers of the children in this study had taken part in an earlier study of fetal growth in white women with singleton pregnancies and known menstrual dates who registered in early pregnancy with participating obstetricians at the Princess Anne Maternity Hospital (Southampton, United Kingdom). Families whose addresses were traceable were contacted for participation in a follow-up study of the impact of fetal growth on physiologic responses to psychosocial stress. The local research ethics committee approved the study, and both parents and children gave written informed consent. A total of 123 children (61 boys and 62 girls with an age range of 7.6–9.8 years) were recorded by a video camera during the whole stress test, and digitized video files were analyzed by 2 observers.

**Procedure**
The children attended a clinical research facility to participate in a version of the Trier Social Stress Test that has been adapted for use in children (TSST-C). For a baseline measure of physiologic parameters, the children were asked to stand in front of a video camera and watch a nonemotional film (Tales of the River Bank) on a screen next to the camera for 5 minutes (nonstress situation). For the TSST-C, the children were asked to stand in front of a microphone and perform an exciting story of their own invention followed by a serial subtraction task for an audience of 3 adult strangers. In both situations, the children were instructed to stand still and to look at the screen where the film was presented (nonstress situation) or at the member of the audience who gave the instructions during the TSST-C. They had 5 minutes to prepare before the stress test, which lasted ~10 minutes. The original TSST-C protocol was modified to reduce task difficulty appropriately for our younger age group, and motivation was increased by offering toys as potential rewards for high performance. At 7 time points during the visit, saliva samples were taken (compare with ref 9), and momentary distress was assessed.

**Measurements**

**Pregnancy Outcomes and Related Measures**
At birth, the infant’s weight was measured using digital scales, and 2 trained field workers recorded neonatal head circumference and crown-heel length using standard techniques; the infant’s gestational age at birth was calculated from the date of the last menstrual period and confirmed by ultrasound. Maternal alcohol or tobacco use during pregnancy was obtained at different occasions by questionnaires. Tobacco use was assessed at the last menstrual period, early pregnancy, and late pregnancy. Alcohol use was assessed at early and late pregnancy. Mother’s socioeconomic class was recorded during the follow-up visit and classified using the United Kingdom National Statistics socioeconomic classification.

**Cortisol and Momentary Distress**
Cortisol concentrations were measured from saliva samples as described in Jones et al. Momentary distress was measured using a visual scale with 5 faces representing variation in affective valence: very positive, rather positive, neutral, rather negative, and very negative. The ends of the scale were explained to the child using examples drawn from their own experience, such as a school or music examination. Scale scores ranged from 0 (very positive) to 4 (very negative).

**Motor Activity**
Both sessions were recorded on videotape, displaying the full body during the complete stress test and the nonstress situation. The tapes were digitized, and the children’s motor activity was quantified using the software program The Observer 5.0 (Noldus Information Systems, Leesburg, USA).
Technology, Wageningen, Netherlands), applying a 1–0 sampling with 15-second sample intervals. Because the children’s arms and trunk were restricted because they were wearing instruments to measure cardiovascular responses (not reported here), the records were confined to motor activity of the feet. The operational definition of motor activity was given as, “at least one foot is lifted or tilted so that at least a part of it loses its contact to the ground.” The motor activity sum score was computed as the sum of the sample intervals during which motor activity occurred. The motor activity rate score was computed as the proportion of all of the sample intervals during which motor activity occurred; this was the principal outcome variable. Each video from the TSST-C was analyzed independently by 2 observers. Because of the very high reliability of the measure, motor activity in the nonstress situation was coded by 1 observer.

**Statistical Analysis**

Differences in the child’s and mother’s characteristics, as well as duration of the TSST-C, motor activity score, and motor activity rate between boys and girls, were tested using 2-tailed t tests for continuous variables and Fisher’s exact tests for categorical variables. Changes in cortisol were tested using paired t tests; changes in momentary distress were tested using Wilcoxon matched-pair signed-rank tests. A random-effects analysis of variance model was used to compute the intraclass correlation coefficient of the motor activity observations as a measure of the interrater reliability. The distribution of the motor activity data was nonnormal, with a high number of 0 scores and a very long right tail of the distribution. Thus, motor activity rates were recoded into 8 equally sized categories with ~15 subjects per group, assuming that this ordinal variable is related to a continuous, latent variable that indicates motor activity. An ordered logit regression model was used to predict the ordinal motor activity rate variable during the TSST-C and the nonstress situation, respectively, by measures of size at birth, and quadratic terms were included where appropriate to test for nonlinear associations. In the case of significant associations, 5 additional models controlled for variables that may confound the associations of size at birth with the children’s behavior. In a first model, the age of the child at the date of the study was included as a predictor. A second model included variables that may affect birth outcomes, namely, parity, mother’s age, and prepregnant BMI, and an additional model included gestational age. A subsequent model included mother’s socioeconomic class to control for adverse environmental influences that are persistent throughout prenatal and postnatal life. Finally, alcohol intake and smoking during pregnancy were added to the model to control for their influence on the associations of size at birth with motor behavior. Because gender differences are common in studies on effects of prenatal adverse factors, all of the models were analyzed for the complete sample and were followed by separate analyses for boys and girls, even if the interaction term with gender was not significant. Stata 9.2 (Stata Corp, College Station, TX) was used for the statistical analysis.

**RESULTS**

Table 1 shows the characteristics of the study sample. The mean gestational age was 280 days, and the mean birth weight was 3.5 kg. Whereas gestational age, birth weight, and age at the date of the stress test did not differ between genders, boys had a higher mean head circumference, length, and ponderal index (PI) at birth than girls. There were no differences between boys and girls concerning the mother’s age at conception, prepregnant BMI, or parity. Twenty-four percent of the mothers smoked during pregnancy; 80% consumed some alcohol while pregnant. However, the median of self-reported number of alcohol units consumed was low, with 0.5 units per week (interquartile range [IQR]: 2.9) in both early and late gestation. There were no differences in mother’s smoking or alcohol consumption between boys and girls.

**Cortisol and Psychological Responses to Stress**

Figure 1 shows the course of mean salivary cortisol and momentary distress values during the children’s visit to the clinical research facility. Both cortisol and distress increased somewhat between measurements 2 and 3 in anticipation of the TSST-C (difference between measurements 2 and 3: cortisol, \( P = .001 \); distress, \( P = .023 \)) and more markedly immediately after the TSST-C stress situation (difference between measurements 3 and 4: cortisol, \( P = .001 \); distress, \( P = .001 \)).

**Motor Activity in Nonstress and Stress Situations**

The interrater reliability of the motor activity sum score for the TSST-C was very high (intraclass correlation coefficient: 0.96). Thus, the mean of the motor activity sum scores recorded by the 2 observers was used in the analyses. The nonstress situation encompassed, on average, 19.6 (SD: 2.6) and the TSST-C, on average, 41.7 (SD: 6.5) sample intervals. The median motor activity sum score for the nonstress situation was 6 (IQR: 8), and for the TSST-C it was 7.5 (IQR: 16.5). The median motor activity rate score, that is, the proportion of all of the sample intervals during which motor activity occurred, was 0.33 (IQR: 0.38) for the nonstress situation and 0.18 (IQR: 0.40) for the TSST-C. The Spearman correlation between the categorized motor activity rate scores during the nonstress and stress situation was an \( r \) of 0.50 (\( P < .001 \)). There were no differences between boys and girls in the mean duration of the situation, the number of sample intervals, the motor activity sum score, or the motor activity rate score for either the nonstress situation or the TSST-C.
Associations of Motor Activity With Birth Weight

The ordered regression analysis for the complete sample showed that the motor activity rates during the nonstress situation and the TSST-C were not associated with birth weight (nonstress: $b = 0.015; \ SE = 0.295; \ P = .961$; TSST-C: $b = 0.366; \ SE = 0.294; \ P = .214$). Testing for gender differences revealed a tendency for differences in the associations for the TSST-C ($P$ for interaction term $= .104$) and no differences in the nonstress situation ($P = .283$). Because the failure to confirm a statistically significant gender difference may be because of limited statistical power and because gender differences are common in studies of the effects of prenatal development, separate analyses for boys and girls were conducted, and the results are shown in Table 2. Motor activity during the nonstress situation was unrelated to birth weight in boys and girls. In contrast, lower birth weight was associated with higher motor activity scores in boys ($P = .048$), whereas no association was observed in girls. When standardized for the latent variable, the significant regression coefficient indicates that for a 1-kg increase in birth weight, the motor activity rate during the stress test in boys is expected to decrease by 0.43 SDs. To illustrate this association, gender-specific quartiles of birth weight were computed, and each cat-

![FIGURE 1](null)

**TABLE 1** Characteristics of Children Participating in the Study and of Their Mothers During Pregnancy

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All ($N = 123$)</th>
<th>Boys ($n = 61$)</th>
<th>Girls ($n = 62$)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child’s characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gestational age, mean (SD), d</td>
<td>279.9 (10.84)</td>
<td>280.0 (11.83)</td>
<td>279.8 (9.87)</td>
<td>.889a</td>
</tr>
<tr>
<td>Birth weight, mean (SD), kg</td>
<td>3.50 (0.51)</td>
<td>3.58 (0.55)</td>
<td>3.43 (0.46)</td>
<td>.116a</td>
</tr>
<tr>
<td>Head circumference at birth, mean (SD), cm</td>
<td>35.10 (1.40)</td>
<td>35.40 (1.59)</td>
<td>34.81 (1.13)</td>
<td>.022a</td>
</tr>
<tr>
<td>Length at birth, mean (SD), cm</td>
<td>50.23 (2.26)</td>
<td>50.76 (2.46)</td>
<td>49.72 (1.94)</td>
<td>.011a</td>
</tr>
<tr>
<td>PI, mean (SD), kg/m²</td>
<td>27.91 (2.19)</td>
<td>27.49 (2.15)</td>
<td>28.33 (2.16)</td>
<td>.034a</td>
</tr>
<tr>
<td>Age at study date, mean (SD), y</td>
<td>8.78 (0.46)</td>
<td>8.83 (0.43)</td>
<td>8.73 (0.48)</td>
<td>.207a</td>
</tr>
<tr>
<td><strong>Mother’s characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at conception, mean (SD), y</td>
<td>28.64 (4.72)</td>
<td>28.20 (4.85)</td>
<td>29.08 (4.59)</td>
<td>.303a</td>
</tr>
<tr>
<td>Prepregnant BMI, mean (SD), kg/m²</td>
<td>23.88 (4.48)</td>
<td>23.72 (4.58)</td>
<td>24.02 (4.40)</td>
<td>.714a</td>
</tr>
<tr>
<td>No. of previous births, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>58 (47)</td>
<td>31 (25)</td>
<td>27 (22)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>45 (37)</td>
<td>23 (19)</td>
<td>22 (18)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>12 (10)</td>
<td>5 (4)</td>
<td>7 (6)</td>
<td></td>
</tr>
<tr>
<td>≥3</td>
<td>8 (7)</td>
<td>2 (2)</td>
<td>6 (5)</td>
<td></td>
</tr>
<tr>
<td>Smoking during pregnancy, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>29 (24)</td>
<td>12 (10)</td>
<td>17 (14)</td>
<td>.396b</td>
</tr>
<tr>
<td>No</td>
<td>94 (76)</td>
<td>49 (40)</td>
<td>45 (36)</td>
<td></td>
</tr>
<tr>
<td>Drinking alcohol during pregnancy, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.655b</td>
</tr>
<tr>
<td>Yes</td>
<td>98 (80)</td>
<td>50 (41)</td>
<td>48 (39)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>25 (20)</td>
<td>11 (9)</td>
<td>14 (11)</td>
<td></td>
</tr>
</tbody>
</table>

* Data are from a $t$ test on equality of means for boys and girls.
* Data are from Fisher’s exact test on differences of proportions in categories according to child’s gender.

**TABLE 2** Results of Ordered Regressions of Motor Activity Rate During a Psychosocial Stress Situation on Birth Weight According to Gender

<table>
<thead>
<tr>
<th>Effect of Birth Weight</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b$</td>
<td>$P$</td>
</tr>
<tr>
<td>Nonstress situation, unadjusted model</td>
<td>0.372</td>
<td>.428</td>
</tr>
<tr>
<td>TSST-C, unadjusted model</td>
<td>0.202</td>
<td>.651</td>
</tr>
<tr>
<td>TSST-C, adjusted model 1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TSST-C, adjusted model 2</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TSST-C, adjusted model 3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TSST-C, adjusted model 4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TSST-C, adjusted model 5</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Similar to a linear regression coefficient, the sign of the regression coefficient $b$ indicates the direction of the association. Model 1 was adjusted for child’s age; model 2 was additionally adjusted for parity, prepregnant BMI, and mother’s age; model 3 was additionally adjusted for gestational age; model 4 was additionally adjusted for social class; and model 5 was additionally adjusted for smoking and drinking alcohol by the mother during pregnancy. — indicates no test performed.

![FIGURE 1](null)

**FIGURE 1**
Geometric mean ($±$SE) of salivary cortisol (dashed line) and mean ($±$SE) of momentary distress (solid line) during the visit of the children at the clinical research facility.
egory’s average motor activity rate during stress for boys is shown in Fig 2. The pattern of mean motor activity rate scores suggests the possibility of a threshold effect, with higher motor activity in children with below-average weight at birth. Formal testing for nonlinearity using a quadratic term was, however, nonsignificant.

As shown in the adjusted models section of Table 2, the association was stable when the following variables were added to the model: child’s age at the study date (adjusted \( P = .034 \)); parity, mother’s prepregnancy BMI, and her age at conception (adjusted \( P = .010 \)); gestational age (adjusted \( P = .024 \)); social status (adjusted \( P = .019 \)); and maternal smoking and alcohol consumption during pregnancy (adjusted \( P = .046 \)).

**Associations of Motor Activity With Head Circumference at Birth**

In boys and girls combined, head circumference at birth was inversely related to motor activity during the TSST-C (\( b = -0.238; SE = 0.115; P = .038 \)) but not to motor activity during the nonstress situation (\( b = -0.046; SE = 0.112; P = .681 \)). The standardized regression coefficient indicates an expected decrease in the motor activity rate score of 0.13 SDs of the latent variable with a 1-cm increase in head circumference. Similar to the results with birth weight, there were no statistically significant differences between boys and girls in the associations, as indicated by the interaction terms in the nonstress situation (\( P = .470 \)) or the TSST-C (\( P = .232 \)). The results of separate analyses for boys and girls are shown in Table 3. There was no association of head circumference at birth with motor activity rate during the nonstress situation in boys, as well as in girls, but a smaller head circumference at birth was associated with higher motor activity rate scores during the TSST-C in boys (\( P = .018 \)). The standardized regression coefficient indicates that for a 1-cm increase in head circumference, the motor activity rate during the stress test in boys is expected to decrease by 0.19 SDs of the latent variable. Figure 3 illustrates this association using gender-specific quartiles of head circumference at birth. The pattern of mean motor activity rate scores suggests the possibility of a threshold effect, with higher motor activity in children with a small head circumference at birth. Formal testing for nonlinearity using a quadratic term was, however, nonsignificant. The adjusted models section of Table 3 documents that this association was stable when poten-
tially confounding variables were controlled: child’s age at the study date (adjusted \( P = .012 \)); parity, mother’s prepregnancy BMI, and her age (adjusted \( P = .004 \)); gestational age (adjusted \( P = .005 \)); social status (adjusted \( P = .005 \)); and maternal smoking and alcohol intake during pregnancy (adjusted \( P = .010 \)).

**Associations of Motor Activity With Length and PI at Birth**

Length at birth was unrelated to motor activity during nonstress (\( P = .691 \)) and stress situations (\( P = .504 \)) in boys and girls combined. Separate analyses for boys and girls also showed nonsignificant associations of length at birth with motor activity during the nonstress (boys: \( b = 0.028; \ P = .762 \); girls: \( b = 0.114; \ P = .312 \)) and stress situations (boys: \( b = -1.121; \ P = .197 \); girls: \( b = 0.085; \ P = .442 \)).

PI at birth showed no association with motor activity during the nonstress situation (\( P = .338 \)) and a weak inverse association during stress (\( b = -0.107; \ P = .136 \)) in boys and girls combined. Separate analyses for boys showed no associations with motor activity during the nonstress situation (\( P = .241 \)) but a significant inverse association of PI at birth with motor activity during the stress situation (\( b = -0.214; \ P = .040 \); no effect of a quadratic term when added to the model); in girls there was no association either during nonstress (\( P = .748 \)) or stress (\( P = .557 \)). The standardized regression coefficient indicates that for a 1-unit increase in PI, the motor activity rate during the stress situation in boys is expected to decrease by 0.11 SDs of the latent variable. Controlling for confounding factors in the model for boys had little effect on these results (adjusted \( P \) values: model 1, \( P = .035 \); model 2, \( P = .043 \); model 3, \( P = .055 \); model 4, \( P = .055 \)); adjustment for maternal smoking and drinking alcohol during pregnancy marginally weakened the effect (model 5, \( P = .099 \)).

A comparison of fully standardized coefficients for the effects in boys showed that the effect of head circumference at birth was strongest (\(-0.30\)), the effects of birth weight (\(-0.24\)) and PI (\(-0.25\)) were similar and slightly smaller, and the (nonsignificant) effect of length at birth was the weakest effect (\(-0.16\)).

**DISCUSSION**

The major finding of this study was an inverse association between size at birth and motor activity stress responses in boys aged 7 to 9 years. Boys who were born with a smaller head circumference, a lower weight, and a lower PI showed higher motor activity during stress, but there was no association of birth size with motor activity during a nonstress situation. The association was stronger for head circumference than for birth weight and PI, and it remained significant after controlling for age of the child, parity, mother’s prepregnant BMI, her age at conception, her social class, and gestational age. The results suggest that gestational age may contribute to the effects observed in our study but that fetal growth restriction rather than preterm birth is the principal factor underlying the associations of birth size with motor activity during stress in this study. This is in contrast with a study in infants that demonstrated a shorter duration of daytime rest in preterm infants compared with term infants.\(^7\) However, our study sample consisted of predominantly term-born children, and our results were confined to motor activity during stress. Therefore, our results are not comparable to these results on activity in preterm infants. Controlling for maternal smoking and drinking alcohol during pregnancy weakened the associations slightly, suggesting that these specific prenatal factors contribute to increased stress-related motor activity in children in our study. Our results suggest that prenatal adversity alters motor activity on exposure to stress.

The results are consistent with the associations of small size at birth, corrected for gestational age, with increased hypothalamic-pituitary-adrenal stress responses\(^8\) and with susceptibility to psychological stress\(^10,11\) in male subjects. Our results thus add to the evidence for the effects of prenatal growth restriction on the biobehavioral stress response later in life. The fully standardized regression coefficients are in the range of \(-0.24\) to \(-0.30\) and demonstrate relatively strong effects compared with coefficients for associations of size at birth with parental ratings of inattention and hyperactivity reported by Lahti et al,\(^4\) which ranged between \(-0.11\) and \(-0.15\). The observation that the associations were confined to boys is consistent with human and animal studies that found marked gender differences in effects of prenatal adverse events on stress reactivity and emotional behavior later in life.\(^7,9,18\)

The observation that size at birth was associated with motor activity during stress, but not during the nonstress situation, suggests different mechanisms for motor activity in these situations. This is corroborated by the moderate correlation between the motor activity rate scores. The marked increase in distress and cortisol during the stress situation further suggests that motor behavior during stress may be related to neuronal circuits that are activated in stressful situations. Whereas dopaminergic innervations of the basal ganglia\(^19\) play a major role in normal motor activity, partly by modulating the influence of prefrontal cortical inputs to striatal neurons,\(^20\) stress triggers a release of dopamine in the striatum and the prefrontal cortex\(^21,22\) and may, thus, influence motor activity. However, our study does not provide conclusive evidence as to the significance of increased motor activity during stress. Increased motor activity may be associated with a number of different states, for example, restlessness, excitement, or a motivation to escape the situation. Because other studies have shown inverse associations of size at birth with stress responses in different domains, we think that in-
increased motor activity during stress may represent a component of the biobehavioral stress response that could have had an adaptive value in an ancient environment (see ref 12). However, our results need to be replicated in future studies, which could use automated movement measurement and should avoid restriction of arm movements.

The finding that head circumference at birth, a good indicator of brain volume,23 was a stronger predictor of motor activity effects than birth weight and PI indicates that the observed functional changes may result from cerebral growth restriction. This is in accordance with recent studies that demonstrated a significantly reduced total brain volume in prenatally growth restricted 15-year-old children24 and a smaller total cerebral volume in patients with ADHD, as well as negative correlations of brain volumes with ADHD symptom severity.25

Because prenatal and postnatal development of the central nervous system is complex,26 and behavioral effects of insults depend on the timing of the event (eg, ref 27), there are a number of potential neurodevelopmental pathways that could mediate this effect. The dopaminergic system is particularly interesting here, because it is involved in motor activity,20 it is activated by stress,21,22 and the prenatal development of dopaminergic neurons seems to be sensitive to growth disruption.28 In addition, there are clear gender differences in dopaminergic neurotransmission,29–32 which could underlie the gender-specific effects observed in our study.

Although we controlled for a number of potential confounding factors, some uncertainty about the key factors remains. For example, although prenatal development may be influenced by parity, mother's age, and prepregnant BMI, these factors did not change the results when included as covariates. We cannot exclude postnatal environmental exposures as the cause of our findings. However, postnatal exposures tend to cluster in low socioeconomic status groups, and inclusion of socioeconomic status as a control variable did not change our results. Furthermore, the relative contribution of genetic and intrauterine environmental factors on fetal growth is unclear. Although birth weight has been shown to reflect the impact of environmental influences on prenatal development,33 it has also been demonstrated to be influenced by genetic factors (eg, refs 34 and 35). On the basis of our data, we cannot draw conclusions about the relative significance of genetic factors, prenatal environmental factors, or their interaction for the associations of size at birth with stress-related motor activity.

Our observations may have clinical implications in relation to hyperactivity and the etiology of ADHD. Our results are consistent with animal studies that demonstrated enhanced locomotor activity in a novel environment after prenatal stress36–38 and with human studies suggesting that prenatal adversity may particularly affect symptoms of ADHD.2–4,39–41 Moreover, motor activity during situations that require sustained attention or mental effort, similar to the TSST-C, has been demonstrated to be enhanced in children with ADHD,44–46 and early adverse events have a permanent impact on relevant neurotransmitter systems.47

CONCLUSIONS

We found that small size at birth was associated with increased motor activity during stress but not during a nonstress situation. The associations were confined to boys and were stable after controlling for potential confounding factors. The anthropometric birth size measures used in this study are only crude measures of prenatal development, and they do not indicate specific factors. However, a permanent effect of the fetal environment on neurodevelopment is a plausible explanation of the effects found in this study, and future studies on potential neurobehavioral mechanisms could provide important insights into the developmental origins of stress-related motor behavior, with potential implications for ADHD.

ACKNOWLEDGMENTS

The study was funded by National Institute of Child Health and Human Development grant 1 R01 HD41107-01. Subjects were drawn from a cohort study funded by WellBeing and the Medical Research Council. Behavioral observations, analysis, and preparation of the article were supported by German Research Foundation (Deutsche Forschungsgemeinschaft) fellowship grant DFG SCHL 1708/1-1 (to Dr Schlotz).

REFERENCES

12. Phillips DIW, Jones A. Fetal programming of autonomic and HPA function: do people who were small babies have enhanced stress responses? J Physiol (Lond). 2006;572:45–50
32. Anderssen SL, Teicher MH. Sex differences in dopamine receptors and their relevance to ADHD. Neurosci Biobehav Rev. 2000;24:137–141
47. Halperin JM, Schulz KP. Revisiting the role of the prefrontal cortex in the pathophysiology of attention-deficit/hyperactivity disorder. Psychol Bull. 2006;132:560–581
Altered Resting Cerebral Blood Flow in Adolescents With in Utero Cocaine Exposure Revealed by Perfusion Functional MRI

Hengyi Rao, PhD*a, Jiongjiong Wang, PhD*a, Joan Giannetta, BAb, Marc Korczykowski, MSa, David Shera, ScDc, Brian B. Avants, PhD*a, James Gee, PhD*a, John A. Detre, MDa, Hallam Hurt, MDh

aDepartment of Radiology and Neurology, Center for Functional Neuroimaging, University of Pennsylvania, Philadelphia, Pennsylvania; bDivisions of Neonatology and cBiostatistics and Epidemiology, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVES. Animal studies have clearly demonstrated the effects of in utero cocaine exposure on neural ontogeny, especially in dopamine-rich areas of cerebral cortex; however, less is known about how in utero cocaine exposure affects longitudinal neurocognitive development of the human brain. We used continuous arterial spin-labeling perfusion functional MRI to measure the effect of in utero cocaine exposure on resting brain function by comparing resting cerebral blood flow of cocaine-exposed adolescents with non–cocaine-exposed control subjects.

PATIENTS AND METHODS. Twenty-four cocaine-exposed adolescents and 25 matched non–cocaine-exposed control subjects underwent structural and perfusion functional MRI during resting states. Direct subtraction, voxel-wise general linear modeling, and region-of-interest analyses were performed on the cerebral blood flow images to compare the resting cerebral blood flow between the 2 groups.

RESULTS. Compared with control subjects, cocaine-exposed adolescents showed significantly reduced global cerebral blood flow. The decrease of cerebral blood flow in cocaine-exposed adolescents was observed mainly in posterior and inferior brain regions, including the occipital cortex and thalamus. After adjusting for global cerebral blood flow, however, a significant increase in relative cerebral blood flow in cocaine-exposed adolescents was found in anterior and superior brain regions, including the prefrontal, cingulate, insular, amygdala, and superior parietal cortex. Furthermore, the functional modulations by in utero cocaine exposure on all of these regions except amygdala cannot be accounted for by the variation in brain anatomy.

CONCLUSIONS. In utero cocaine exposure may reduce global cerebral blood flow, and this reduction may persist into adolescence. The relative increase of cerebral blood flow in anterior and superior brain regions in cocaine-exposed adolescent participants suggests that compensatory mechanisms for reduced global cerebral blood
flow may develop during neural ontogeny. Arterial spin-labeling perfusion MRI may be a valuable tool for investigating the long-term effects of in utero drug exposure.

IN UTERO COCAINE exposure (IUCE) and its potential cognitive and psychological sequelae have been major health-related concerns since the late 1980s in the United States. Although theories and models from animal experiments have clearly demonstrated some alterations and deficits of cortical development induced by IUCE, especially in the dopamine-rich areas of cerebral cortex, the effect of IUCE on human subjects remains unclear. On one hand, evidence from previous studies has shown different degrees of harmful IUCE effects in the newborn, such as premature birth, birth asphyxia, brain hemorrhage, growth retardation, and neurobehavioral deficits. On the other hand, a review of IUCE effect in early childhood failed to find any consistent association between IUCE and physical growth or cognitive developmental deficits. Currently, there is still no clear answer regarding the extent to which IUCE may have an enduring impact on the neurocognitive development of children. The inconsistency of the IUCE effect on human subjects may be because of the complex methodologic issues involved in conducting longitudinal studies with the diverse human population. The unavoidable imprecision in ascertaining the gestational timing and dose of cocaine to which the fetus was exposed, the occurrence of concurrent substance use other than cocaine, the socioeconomic status and family mental health history, some limited sample sizes, and the lack of long-term follow-up may all contribute to the inconsistency.

Recent advances in brain imaging and cognitive neuroscience have led to growing interest in integrating neuroimaging methods and neurocognitive framework into the study of IUCE effects. Using different imaging techniques, such as functional or structural MRI, diffusion tensor MRI, magnetic resonance spectroscopy, electroencephalography, or event related potential, previous studies have provided emerging evidence suggesting that IUCE may affect the frontal lobe, which may be reflected in neurocognitive impairments in attentional and arousal regulatory systems, as well as executive function.

Longitudinal studies of cortical development from childhood through adulthood in healthy normal brains have demonstrated that maturation of high-order association cortices occurs later than low-order sensory cortices, such as the maturational changes of the prefrontal cortex during late adolescence. In addition, animal studies on nonhuman primates suggest that the behavioral effects of IUCE might not be manifested until adulthood. However, to date, few studies have investigated the impact of IUCE on human brain activity during adolescence. Our group has followed a cohort of in utero cocaine-exposed and nonexposed participants since their birth in the late 1980s or early 1990s. These participants, who are now entering early-middle adolescence, provide an opportunity to investigate the effects of IUCE on neurocognitive function in the adolescent brain.

To this end, the present study used arterial spin-labeling (ASL) perfusion functional MRI (fMRI) and measured resting cerebral blood flow in 2 groups of adolescents, with and without IUCE, with the aim of noninvasively exploring the effects of IUCE on resting cerebral activation patterns in the adolescent brain. ASL perfusion MRI offers absolute quantification of cerebral blood flow (CBF; ie, milliliters of blood per 100 g of tissue per minute) that is normally coupled with neural activity by using magnetically labeled arterial blood water as an endogenous tracer. Previous studies have shown excellent reproducibility over long-term time periods and less between-subject variability of ASL perfusion compared with blood oxygen-dependent level (BOLD) fMRI. ASL may be particularly advantageous for pediatric neuroimaging, because it is entirely noninvasive and provides improved image quality compared with adult images because of several physiologic properties of a child brain. We were particularly interested in the effect of IUCE on the frontal lobe and the reciprocally connected limbic structures. Therefore, we defined a priori regions of interest (ROIs) of frontal lobe and the limbic structures including the cingulate cortex, caudate, insula, and amygdala. The occipital lobe and thalamus were also included as the lower-order sensory processing regions for comparison with the frontal lobe. To explore any possible association between the influences of IUCE on resting brain CBF and on brain anatomy, perfusion fMRI was combined with optimized voxel-based morphometry (VBM), a quantitative morphometrical analysis of structural MRI to compare the gray matter volume between groups.

METHODS

Participants
A total of 49 adolescent participants, including 25 cocaine-exposed adolescent (COC) participants and 24 non–cocaine-exposed control subject (CON) participants, were recruited from a cohort of exposed and nonexposed participants who have been followed since their birth (1989–1992). Full details of enrollment have been reported previously. Written consent was obtained according to institutional review board approval from the Children’s Hospital of Philadelphia. Subjects selected for MRI met the following criteria: they were right handed, had no metal appliances, and were on no medications. Because of concerns for possible confounding effects of gender and IQ, we further selected exposed
and nonexposed children by gender and by group quartiles of 4-year Wechsler Preschool and Primary Scale of Intelligence-Revised scores. We did not perform drug screens before the fMRI scan. However, all of the subjects in the cohort have drug screens at the time of the annual visit, and there was no difference between the COC and CON groups. All of the participants were born at a single inner-city hospital to mothers of low socioeconomic status. The COC cohort was heavily cocaine exposed during gestation with 24 of 25 having been exposed in all 3 trimesters and 19 of 25 mothers of the COC cohort having used frequently (once or more per week; median: 117 days of exposure and a minimum of 2 trimesters of pregnancy); conversely, the mothers in the CON cohort denied cocaine use, and both mother and child had urine samples (peripartum and natal, respectively) that were negative for cocaine metabolites.

Among the COC group, 8 mothers reported using cigarettes, alcohol, and marijuana; 7 mothers reported using cigarettes and alcohol; 5 mothers reported using cigarettes and marijuana; 4 mothers reported using cigarettes; and 1 mother reported using marijuana; in the CON group, only 1 mother reported using cigarettes, and no mothers reported using alcohol or marijuana. The numbers of mothers using cigarettes, alcohol, or marijuana in the 2 groups are listed in Table 1. Children were discharged before the scan could be performed. Participants also had a neurologic examination at 6.5 years of age by a developmental pediatrician masked to group status who used a standard neurologic examination (cranial nerves, optic fundus, deep-tendon reflexes, tone, gross motor strength, range of motion, gait, plantar responses, sensation, and an articulation screen) and an examination for soft signs (finger to nose, rapid pronation-supination, heel-to-toe walking, balance on 1 foot with eyes closed, and hop) to evaluate the children.

### Imaging Acquisition

A continuous ASL technique was conducted on a Siemens 3.0-T Trio whole-body scanner (Siemens AG, Erlangen, Germany) using a standard transmit/receive head coil for perfusion fMRI scans. ASL was implemented with a 0.16-g/cm gradient and 22.5-mg radiofrequency irradiation applied 8 cm beneath the center of the acquired slices. Control/labeling was interleaved using an amplitude-modulated version of the labeling pulse based on a sinusoid function. The tagging/control duration was 1.6 seconds. Interleaved images with and without labeling were acquired using a gradient echoplanar imaging sequence. A delay of 1.2 seconds was inserted between the end of the labeling pulse and image acquisition to reduce transit artifact. Acquisition parameters were as follows: field of vision, 22 × 22 cm²; matrix, 64 × 64; repetition time (TR)/echo time, 4000 milliseconds/17 milliseconds; flip angle, 90°. The resting perfusion scanning protocol lasted 320 seconds with 80 acquisitions. Sixteen slices (6-mm thickness

### TABLE 1 Study Demographics of the 49 Participants

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Cocaine-Exposed</th>
<th>Control</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of participants</td>
<td>25</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Gender, female, n/N (%)</td>
<td>11/25 (44)</td>
<td>13/24 (54)</td>
<td>.57</td>
</tr>
<tr>
<td>Age at testing, mean ± SD, y</td>
<td>14.4 ± 1.0</td>
<td>13.9 ± 0.9</td>
<td>.06</td>
</tr>
<tr>
<td>Race, black, n/N (%)</td>
<td>24/25 (96)</td>
<td>24/24 (100)</td>
<td>.32</td>
</tr>
<tr>
<td>Urine test positive for cocaine at delivery, n/N (%)</td>
<td>18/22 (82)</td>
<td>0/24 (0)</td>
<td>.00</td>
</tr>
<tr>
<td>Cocaine use in pregnancy, median (range), d</td>
<td>117 (7–273)</td>
<td>0 (0)</td>
<td>.00</td>
</tr>
<tr>
<td>Marijuana use in pregnancy, n/N (%)</td>
<td>14/24 (58)</td>
<td>0 (0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Cigarette use in pregnancy, n/N (%)</td>
<td>24/25 (96)</td>
<td>1/24 (4)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Alcohol use in pregnancy, n/N (%)</td>
<td>15/25 (60)</td>
<td>0 (0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Birth weight &lt; 10th percentile, n/N (%)</td>
<td>0/25 (0)</td>
<td>1/24 (4)</td>
<td>.49</td>
</tr>
<tr>
<td>Head circumference &lt; 10th percentile, n/N (%)</td>
<td>3/25 (12)</td>
<td>2/24 (8)</td>
<td>.67</td>
</tr>
<tr>
<td>Gestational age, mean ± SD (range), wk</td>
<td>38.0 ± 2.4 (34–42)</td>
<td>39.1 ± 2.0 (34–42)</td>
<td>.09</td>
</tr>
<tr>
<td>Apgar score at 5 min, median (range)</td>
<td>9 (6–10)</td>
<td>9 (8–9)</td>
<td>.59</td>
</tr>
<tr>
<td>Cranial ultrasound abnormality, n/N</td>
<td>0/25</td>
<td>0/21</td>
<td>.00</td>
</tr>
<tr>
<td>Any neurologic abnormality at 6.5 y, n/N (%)</td>
<td>3/20 (15)</td>
<td>2/19 (10)</td>
<td>.68</td>
</tr>
<tr>
<td>&gt;1 neurologic abnormality at 6.5 y, n/N (%)</td>
<td>1/20 (5)</td>
<td>1/19 (5)</td>
<td>.97</td>
</tr>
<tr>
<td>Full-scale IQ at 6 y, mean ± SD</td>
<td>86.6 ± 13.5</td>
<td>89.1 ± 10.2</td>
<td>.52</td>
</tr>
</tbody>
</table>
with 1.5-mm gap) were acquired from inferior to superior in sequential order. Before the perfusion scan, high-resolution anatomic images were obtained by a three-dimensional magnetization-prepared rapid acquisition of gradient echo sequence with TR/inversion time/echo time at 1620 milliseconds/950 milliseconds/3 milliseconds, flip angle at 15°, 160 contiguous slices of 1-mm thickness, field of vision at 192 × 256 mm², and matrix at 192 × 256. The total length of scan time lasted ~1 hour, including the perfusion scan, anatomic scan, and other scans for BOLD imaging and diffusion tensor imaging (data not reported here).

Functional Imaging Data Analysis
Functional and structural MRI data processing and analysis were conducted primarily with the statistical parametric mapping software (SPM2 [Wellcome Department of Cognitive Neurology, United Kingdom], implemented in Matlab 6 [Math Works, Natick, MA]), with some modifications for perfusion analysis (see http://afni. nhs.ucl.ac.uk/spm/software.htm).

For each subject, functional images were first realigned to correct for head motion and coregistered with the anatomic image. Perfusion-weighted image series were then generated by pairwise subtraction of the label and control images, followed by conversion to absolute CBF image series based on a single-compartment continuous arterial spin-labeling perfusion model. Thus, the resulting CBF data sets contained 40 acquisitions with an effective TR of 8 seconds. One mean CBF image was generated for each individual subject, normalized to a 2 × 2 × 2-mm³ subject-based Montreal Neurologic Institute template generated from the optimized VBM protocol described in previous studies. The VBM approach provides a sophisticated automated method to measure gray matter volume differences between 2 groups. A study-specific T1 brain template was created for spatial normalization, and study-specific probability maps were created to optimize the segmentation of each subject’s image. The customized T1 template, gray matter, white matter, and cerebrospinal fluid images were used for the optimized VBM procedure. The spatially normalized segments of each subject’s gray matter images were modulated for volume analysis and smoothed in space with a three-dimensional, 8-mm full-width-at-half-maximum Gaussian kernel. Two-sample t tests were performed on the global volumes of gray matter, white matter, and cerebrospinal fluid. The individual modulated gray matter images were entered into the whole brain voxel-wise general linear model with 3 covariates to account for the age, gender, and total gray matter volume variability. Areas that showed significant difference between the 2 groups were identified for the mapwise significance level of the FDR-corrected P value <.05 and cluster size >100 voxels. The gray matter volumes of the above ROIs were calculated. To investigate any possible association between the influences of IUCE on brain anatomy and on resting CBF, multivariate regression analyses were performed on regional absolute CBF in each ROI, using IUCE, total gray matter volume, and regional gray matter volume in each ROI as the independent covariates.

RESULTS
The study demographics of all 49 of the participants are shown in Table 1. The COC and CON groups were similar (all P values at >.05) except for poly-substance exposure in COC (all P values at <.001). The quantitative resting CBF images averaged from the COC group and CON group, as well as the CBF differences, are shown in Fig 1. Both resting CBF images visualized the perfusion signal in all of the brain regions with good sensitivity and illustrated clear contrast between gray and white matter in the perfusion intensity. The global
CBF intensities were significantly lower for the COC group than the CON group (10.1% decrease; *P* = .047; Fig 3B). The direct subtraction of the CBF image of the COC group from that of the CON group revealed robust CBF decreases in widespread posterior and inferior brain regions (Fig 1C). These regions included the occipital cortex and thalamus and extended to the cerebellum, fusiform, posterior hippocampus, and inferior temporal cortex. Sparse CBF increases were observed in small anterior frontal, cingulate, and parietal regions.

The results from the voxel-wise analysis are shown in Fig 2 and listed in Table 2. Without correction of the global CBF difference, significant CBF decreases were found in a large posterior and inferior brain region that were similar to the finding from direct subtraction (Fig 2A). However, with correction of global CBF, significant relative CBF increases were found in multiple anterior and superior brain regions, including bilateral medial and dorsal frontal cortex, insula-putamen, cingulate, and parietal cortices. The regional CBF results for the
ROI analyses are shown in Fig 3. Significant absolute CBF (Fig 3B) decreases in the COC group were observed in the occipital lobe (19.8% decrease; \( P = .004 \)) and the thalamus (21.1% decrease; \( P = .003 \)), whereas significant relative CBF (Fig 3C) increases were seen in the frontal lobe (9.3% increase; \( P = .001 \)), cingulate cortex (12.2% increase; \( P < .001 \)), insula (10.2% increase; \( P < .001 \)), and amygdala (12.2% increase; \( P = .004 \)). These ROI results confirmed the voxel-wise results. However, no absolute or relative CBF difference was found in the caudate (both \( P > .05 \)).

From morphometrical analyses, there were no group differences in the total volumes of gray matter, white matter, or cerebrospinal fluid (all \( P > .1 \); data not shown). The voxel-wise comparison revealed no significant differences between the gray matter volumes of the 2 groups when using the whole-brain mapwise corrected threshold. The ROI analyses revealed no group differences except that the relative regional gray matter volume (percentage of total gray matter volume) was significantly decreased in caudate (5.1% decrease; \( P = .02 \)), and the absolute and relative regional gray matter volumes were both significantly increased in amygdala (6.2% and 7.0% increase; both \( P < .02 \)) in the COC group. With the regional gray matter volume and total gray matter volume included in the multiple regression analyses of the CBF data, the IUCE effect was still significant for absolute resting CBF reduction in the thalamus and occipital lobe (both \( P < .01 \)), as well as for relative resting CBF increase in the frontal lobe, cingulate cortex, insula, thalamus, and occipital lobe (all \( P < .05 \); Table 3).

**DISCUSSION**

Capitalizing on the several advantages of perfusion fMRI and our longitudinal cohort followed since birth, the present study measured resting brain CBF in a relatively large group of adolescent participants with and without IUCE and suggests several important findings with respect to the effects of IUCE on the developing human brain during adolescence.

The first finding of the present study is the reduced global CBF in prenatally cocaine-exposed participants compared with nonexposed participants. Reduction of cerebral glucose metabolism and CBF in the human brain is a well-established neurologic consequence of cocaine abuse in adults.\(^{34-38} \) Cocaine use can lead to both a global CBF decrease and regional hypoperfusion. To quantify the effects of cocaine use on global CBF, Wallace et al\(^{39} \) used single photon emission computed tomography and reported an \( \sim 30\% \) decrease in absolute whole-brain CBF at the time of peak cocaine subjective effects. Similarly, Johnson et al\(^{35,36} \) also used single photon emission computed tomography and reported an \( \sim 8\% \) to 10% decrease in whole-brain blood flow for intravenous cocaine administration compared with recently abstinent cocaine-dependent subjects. In addition, Gollub et al\(^{40} \) reported a 14% decrease in CBF to cortical gray matter 15 to 30 minutes after infusion of cocaine using flow-sensitive alternating inversion recovery MRI. In the present study, IUCE was associated with a 10% reduction in global CBF, consistent with these values reported previously. A previous study with adult subjects has demonstrated that cocaine has potent effects of cerebral vasoconstriction, and chronic exposure may alter cerebrovascular reactivity and permanently de-
crease CBF. Reduced global CBF in subjects with IUCE suggests that chronic exposure in utero may have similar effects on the fetus, and these effects may persist into adolescence.

The second finding of this study is the altered distribution of resting CBF in participants with IUCE compared with control subjects. The direct subtraction of quantitative CBF, the voxel-wise general linear modeling, and the ROI analyses all consistently showed a reduction in absolute CBF, primarily in posterior and inferior brain regions, including the occipital lobe, thalamus, and posterior temporal lobe. However, frontal, cingulate, insula, and caudate regions showed little absolute CBF differences. The occipital lobe is the cortical center for visual processing, and the thalamus is the subcortical center for receiving and projecting visual and other sensory signals to cortex. Specifically, 1 previous study has demonstrated systematically underdeveloped gray matter cellular structure in the occipital lobe of primates with IUCE. In adult human subjects with chronic cocaine abuse, Lee et al reported an enhanced BOLD signal in occipital regions in response to photic stimulation, which may reflect inefficient neuronal processing of visual information. The present results are in line with and further extend these findings in terms of IUCE-reduced resting brain function of visual occipital processing in adolescents.

Third, after correction for global CBF differences, the relative CBF was significantly higher in frontal, cingulate, insula, amygdala, and parietal regions in participants with IUCE, that is, relatively more CBF was distributed to the frontal and superior brain regions than the posterior and inferior brain regions. These frontal, cingulate, and parietal regions serve as the neural substrates mediating attention and arousal regulation, which are higher-order association cortices, the development of which is far from complete by childhood. The changes in relative CBF in these regions are consistent with the view that IUCE-induced changes may affect attention processing. That CBF relatively increased in the late-maturing frontal regions but decreased in the early maturing occipital regions in the cohort with IUCE may reflect the compensatory mechanisms involved in the brain development.

Interestingly, we observed significant relative CBF increases in amygdala and insula but not in caudate. These brain structures are both critical for processing emotion and for linking emotion to behavior. The amygdala and insula are key components of the neural network that mediates emotion processing and emotional reactivity to environmental stimuli. The changes in relative CBF in these regions also suggest that adolescents with IUCE may have altered emotional processing and reactivity. Furthermore, the increased CBF in the frontal and parietal regions observed in this study may reflect compensatory mechanisms to compensate for reduced functional connectivity in the visual and attentional networks. These findings highlight the importance of considering the psychosocial aspects of IUCE and the potential for long-term neurodevelopmental sequelae.

### Table 2: Voxelwise GLM Revealed Absolute and Relative Resting CBF Differences Between the COC and CON Groups

<table>
<thead>
<tr>
<th>Brain Regions</th>
<th>MNI Coordinates</th>
<th>z Score</th>
<th>Corrected P</th>
<th>Cluster Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Absolute CBF: COC &lt; CON</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilateral occipital, fusiform, cerebellum, thalamus, inferior temporal</td>
<td>4</td>
<td>-52</td>
<td>12</td>
<td>4.32</td>
</tr>
<tr>
<td>6</td>
<td>-22</td>
<td>-16</td>
<td>4.06</td>
<td>.029</td>
</tr>
<tr>
<td><strong>Relative CBF: COC &gt; CON</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left insula-putamen</td>
<td>-42</td>
<td>-10</td>
<td>-6</td>
<td>5.03</td>
</tr>
<tr>
<td>Left parietal</td>
<td>-58</td>
<td>-30</td>
<td>20</td>
<td>4.06</td>
</tr>
<tr>
<td>Right parietal</td>
<td>42</td>
<td>-50</td>
<td>36</td>
<td>4.81</td>
</tr>
<tr>
<td>Right frontal</td>
<td>34</td>
<td>40</td>
<td>40</td>
<td>4.03</td>
</tr>
<tr>
<td>Right insula-putamen</td>
<td>36</td>
<td>-14</td>
<td>2</td>
<td>3.53</td>
</tr>
<tr>
<td>Right angular</td>
<td>-36</td>
<td>-76</td>
<td>36</td>
<td>3.44</td>
</tr>
<tr>
<td>Bilateral posterior cingulate</td>
<td>2</td>
<td>-22</td>
<td>30</td>
<td>4.31</td>
</tr>
<tr>
<td>Bilateral medial frontal</td>
<td>12</td>
<td>46</td>
<td>0</td>
<td>4.06</td>
</tr>
<tr>
<td>Bilateral anterior Cingulate</td>
<td>0</td>
<td>20</td>
<td>26</td>
<td>3.40</td>
</tr>
</tbody>
</table>

Threshold was set as an FDR-corrected P value of <.05 and cluster size >100 voxels. MNI indicates Montreal Neurological Institute.

### Table 3: Multiple Regression Analyses of the Absolute and Relative Resting Regional CBF Values in the ROIs With the Regional Gray Matter Volume and Total Gray Matter Volume as Confounding Covariates

<table>
<thead>
<tr>
<th>ROI</th>
<th>Frontal Lobe</th>
<th>Cingulate Cortex</th>
<th>Insula</th>
<th>Amygdala</th>
<th>Caudate</th>
<th>Occipital Lobe</th>
<th>Thalamus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Absolute CBF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explained variance, %</td>
<td>0.005</td>
<td>0.180</td>
<td>0.450</td>
<td>1.700</td>
<td>0.490</td>
<td>16.300</td>
<td>16.800</td>
</tr>
<tr>
<td>P</td>
<td>.99</td>
<td>.78</td>
<td>.66</td>
<td>.46</td>
<td>.65</td>
<td>.004&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.006&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Relative CBF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explained variance, %</td>
<td>22.90</td>
<td>26.40</td>
<td>20.30</td>
<td>6.50</td>
<td>0.07</td>
<td>11.60</td>
<td>13.50</td>
</tr>
<tr>
<td>P</td>
<td>.001&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.001&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.11</td>
<td>.08</td>
<td>.02&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.02&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

The variance explained by the COC effect is shown.

<sup>a</sup> P < .005
<sup>b</sup> P < .05.
dala and insula are well known for their critical roles in the processing of aversive and unpleasant affects, such as fear, disgust, and pain. The observed CBF enhancement in amygdala and insula is consistent with the evidence showing the IUCE effect on negative affects. The caudate is well known to be involved in the processing of positive affects, such as reward. There were no CBF changes in caudate may provide impetus for future studies to examine the neurologic basis associated with the possible different alterations in negative and positive affects in children with IUCE.

In addition to investigating the resting CBF difference between the 2 groups, the present study also used morphometrical analysis to explore the possible effects of IUCE on the brain anatomy of adolescent participants. The results demonstrated no significant effect of IUCE on the whole brain or regional gray matter volume, with the exceptions of increases in absolute and relative amygdala volume and decrease in relative caudate volume revealed by ROI analysis. The less significant anatomic differences between exposed and nonexposed participants suggest either a dissociation of the IUCE effect on brain function or brain anatomy or a higher sensitivity of perfusion fMRI than morphometrical analysis. Furthermore, the relative volume difference observed in the caudate was confirmed by the voxel-wise analysis of the same data using a more sensitive method, symmetric diffeomorphisms, for brain structural analysis. When including the regional gray matter volume and total gray matter volume in the multivariate regression analyses, the IUCE effect was still significant for absolute CBF changes in the occipital cortex and thalamus, as well as for relative CBF changes in the frontal cortex, insula, and cingulate cortex. These findings suggest that CBF alterations in these regions in participants with IUCE cannot be accounted for by variations in gross brain anatomy. However, the IUCE effect on the absolute and relative CBF in amygdala was not significant when adjusting amygdala anatomic differences, suggesting that IUCE may affect the structure and function of amygdala in a similar way, or structural change of amygdala may be the reason for increased amygdala CBF. These results raised interesting issues for further research regarding the relationship and the temporal order of the IUCE effects on brain function and anatomy.

There were several limitations in our study. First, drug-abusing adults rarely use cocaine in isolation and usually combine cocaine with other potential neurotoxins, such as alcohol, tobacco, and marijuana; thus, the interpretation of the effect of drugs in humans often is confounded. In the present study, the COC cohort was also exposed to multiple drugs, and there were robust differences of poly-drug exposure between the COC and CON cohorts; thus, the present findings might be confounded by concurrent exposure of tobacco, alcohol, and marijuana in mothers. To minimize this confounding effect, we separated the COC group by each kind of drug exposure and performed additional comparisons between the COC and CON groups by only exposed subjects without a given type of drug exposure (eg, to minimize the effect of alcohol exposure, we compared the COC and CON groups using only subjects without alcohol exposure). We also performed comparisons between the subjects with and without a given type of drug exposure within the COC group (eg, we compared the COC subjects without alcohol exposure with the COC subjects with alcohol exposure). Results from these additional analyses showed similar patterns, suggesting that the altered resting CBF patterns observed in adolescents were associated with IUCE rather than poly-drug use. Future studies with better control of confounding variables in a larger cohort will be needed to elucidate the independent effect of IUCE on brain function and development.

Second, the subjects were pubescent or peripubescent, thus, hormonal changes, especially in female subjects, might interact with the IUCE. However, the analyses including only male subjects yielded similar results, suggesting that the present effects are unlikely to be related to the pubescent hormonal changes.

Finally, all of the subjects were from families of low socioeconomic status, and their IQ scores were mildly low. Whether IUCE affects children of high socioeconomic status and normal IQ the same way and how environmental enrichment and parental nurturance interact with IUCE during brain development remain open to future studies.

CONCLUSIONS
We used ASL perfusion fMRI and noninvasively quantified resting CBF in adolescent participants with and without IUCE, demonstrating that IUCE reduces global CBF and alters CBF distribution in the adolescent brain. To our knowledge this is the first study to investigate the impact of IUCE on resting brain function in adolescence. The observed global CBF reduction suggests that the IUCE effect on blood flow may persist until adolescence or young adulthood. In light of reduced global CBF, the apparent increase in the distribution of CBF to more frontal and superior brain regions in exposed participants suggests that compensatory mechanisms may be involved during neurodevelopment. Whether and how these alterations in CBF will be reflected in behavior changes in adolescents and young adults with IUCE are the focus of an ongoing investigation.

Using neuroimaging techniques to image brain structure, function, and metabolites in longitudinally followed exposed and nonexposed control children has emerged as one of the major research directions in the study of prenatal drug exposure. This current study and previous reports show the promise of neuroimaging studies to enhance our understanding of how ma-
ternal drug abuse affects neurobehavior in their offspring. Perfusion fMRI will be a valuable tool for imaging the long-term effects of drug use during pregnancy in neurodevelopment, given its excellent stability over time.

ACKNOWLEDGMENTS

This research was supported in part by National Institutes of Health Human Brain Project MH072576 and HD049893, National Science Foundation grant BCS-0224007, the Thrasher Research Fund, National Institute on Drug Abuse grant DA14129, and National Institute of Child Health and Human Development grant MRDDRC-HD26979.

REFERENCES

38. London ED, Bonson KR, Ernst M, Grant S. Brain imaging
Scoring Evaluation of the Natural Course of Mucopolysaccharidosis Type IIIA (Sanfilippo Syndrome Type A)

Ann Meyer, MD\(^{a}\), Kai Kossow, PhD\(^{b}\), Andreas Gal, MD\(^{c}\), Chris Mühlausen, MD\(^{a}\), Kurt Ullrich, MD\(^{a}\), Thomas Braulke, PhD\(^{a}\), Nicole Muschol, MD\(^{a,c}\)

\(^{a}\)Department of Pediatrics, Metabolic Center, \(^{b}\)Center of Psychosocial Medicine, Institute and Policlinics of Medical Psychology, and \(^{c}\)Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Mucopolysaccharidosis types IIIA through IIDD (Sanfilippo syndrome) are caused by deficiencies of enzymes involved in the degradation of heparan sulfate. The onset and severity of the disease are highly variable. The purpose of this study was to describe the natural course of mucopolysaccharidosis type IIIA in a large cohort of patients.

PATIENTS AND METHODS. The natural course of mucopolysaccharidosis type IIIA was assessed in 71 patients by using a detailed questionnaire and a 4-point scoring system and compared with the course of the disease in 14 patients with mucopolysaccharidosis type IIIB and 4 patients with mucopolysaccharidosis type IIIC.

RESULTS. In the cohort of patients with mucopolysaccharidosis type IIIA, first symptoms of disease were observed, on average, at 7 months of age. Speech and motor development were delayed in 66.2% and 33.9% of patients, respectively. The median age at diagnosis was 4.5 years. The onset of regression in speech, motor, and cognitive function was observed at an average age of 3.3 years. The loss of all 3 of the assessed abilities was observed at an average age of 12.5 years. Speech was lost before motor and cognitive functions. In a small group of patients who were >12.5 years of age (9.9%), speech, motor, and cognitive skills were partially preserved up to a maximum age of 23.8 years.

CONCLUSIONS. To our knowledge, this is the first systematic and comprehensive study on the natural course of mucopolysaccharidosis type IIIA. The 4-point scoring system may be used to classify patients into groups with a rapid or slower course of the disease. This may have an important impact on parental counseling as well as therapeutic interventions.
**Mucopolysaccharidosis (MPS) Type III (Sanfilippo Syndrome)** is a genetically and clinically heterogeneous group of diseases caused by the deficiency of 1 of 4 enzymes (defining subtypes A–D) involved in the degradation of heparan sulfate. MPS IIIA arises when N-sulfoglucosamine sulfohydrolase activity (sulfamidase, EC 3.10.1.1) is lost. MPS IIIB and C are caused by defective α-N-acetylgalcosaminidase (EC 3.2.1.50) and heparan-α-glucosaminidase (EC 2.1.3.78), respectively. The activity of N-acetylgalactosamine-6-sulfatase (EC 3.1.6.14) is deficient in patients with MPS IIID. All of the MPS III subtypes are inherited in an autosomal recessive manner.1 The incidence of MPS III in Germany has been estimated as 1 in 63,700 births.2 MPS IIIA is the most common subtype in Northern Europe, whereas MPS IIIB is more prevalent in Southern Europe.3,4

The clinical features of the different enzyme deficiencies are difficult to distinguish. Patients with MPS III have been reported to be normal at birth and develop normally during the first year of life. Developmental delay becomes apparent in early childhood, and children may exhibit behavioral abnormalities, sleep disturbances, and speech delay. Somatic features are often mild and variable. Patients may show coarse facial features, mild skeletal dysostosis multiplex, and hepatosplenomegaly. Mental retardation becomes obvious over time, and progressive neurologic degeneration occurs. Patients generally die within the second decade of life; however, survival into the third or fourth decade has also been reported.5 There may be differences in the severity of disease in the different subtypes. Whereas adult onset of dementia has been reported in MPS IIIB patients,5–7 dementia is apparent in the majority of patients with MPS IIIA by the age of 6 years.8 Progression of the disease is thought to be more rapid in MPS IIIA than in MPS IIIB and C. However, the clinical phenotype of patients with the same subtype of MPS III, even in siblings, is highly variable.9–13 The severity of MPS IIIC is reported to lie between that of MPS IIIA and B.14 MPS IIID is very rare and also heterogeneous.15,16 The variability in the clinical phenotype of patients with MPS III is presumed to be because of variations in residual enzyme activity, caused by the different homozygous and compound heterozygous mutations.3,17–19

In this study, the natural course of MPS IIIA was investigated in a group of 71 patients and compared, where possible, with the progression of the disease in 14 patients with MPS IIIB and 4 patients with MPS IIIC. The study was based on a questionnaire and a 4-point scoring system (FPSS). This study describes the natural course of the disease in a large cohort of German patients with MPS III.
clinical and social care, kindergarten and school attendance, interests and hobbies, developmental regression, and death. The progression of the disease was tracked using a modified scoring system described previously for other neurodegenerative disorders of childhood.\textsuperscript{20,21} The FPSS assessed motor function, speech abilities, and cognitive function, retrospectively, over the course of disease in 3- to 6-month intervals. Scoring assessments were made as follows: 3 points for normal function, 2 points at the beginning of regression, 1 point where regression had progressed to a severe level, and 0 points where function was lost. Detailed information on the scoring method used in this study is given in Table 1. The total disability score (TDS) is the average of the motor function, speech, and cognitive function scores. The interviews and scoring assessments were all conducted by the same interviewer. The detailed questionnaire is available on request from the authors. The study was approved by the medical ethics committee of the Ärztekammer Hamburg.

MPS III Subtype Classification
MPS III subtype was determined through sulfamidase, α-N-acetylgalcosaminidase, and heparan-α-glucosaminide N-acetyltransferase activity analysis in either white blood cells or fibroblasts. Quantitative glycosaminoglycan analysis was performed using the dimethylmethyl blue method.\textsuperscript{22} Two-dimensional electrophoresis of urinary glycosaminoglycans was conducted as described previously.\textsuperscript{21}

Data Handling and Analysis
To protect patient privacy, each patient received a study number, which was used in all of the analyses. Personal data were separated from the questionnaire. All of the analyses were performed by using SPSS 12.0 for Windows (SPSS Inc, Chicago, IL).

<table>
<thead>
<tr>
<th>Function</th>
<th>Performance</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor function\textsuperscript{a}</td>
<td>Normal walking</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Clumsy walking</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Aided walking</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Wheel chair/immobile</td>
<td>0</td>
</tr>
<tr>
<td>Speech abilities\textsuperscript{a}</td>
<td>Normal speech</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Impairment of speech</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Speech difficult to understand</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Loss of speech</td>
<td>0</td>
</tr>
<tr>
<td>Cognitive function</td>
<td>Normal cognitive function</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Deterioration of cognitive function</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Loss of interest in environment</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Unresponsiveness</td>
<td>0</td>
</tr>
</tbody>
</table>

\textsuperscript{a} In some patients, motor and speech development were never normal. In these case subjects, the scoring started at a score of 2.

RESULTS
A total of 89 patients with MPS III (51.7% males and 48.3% females) from 82 families were included in the study; this total was composed of 71 patients (79.8%) with subtype A, 14 patients (15.7%) with subtype B, and 4 patients (4.5%) with subtype C. No patients with MPS IIID took part in the study. Ten patients (11.2%) had died at the time of data collection. The age and gender distribution of the study population at the time of data collection is shown in Table 2. Because of the low number of patients with subtypes B (\(n = 14\)) and C (\(n = 4\)), comparative statistical analysis of the data was not possible. Similarities in the clinical features, course, and progression of the disease, however, were mentioned, where relevant. The average age of patients with MPS IIIA at the time of data collection was 13.2 years (SD: 6.8 years [range: 1.8 –32.8 years]).

Pregnancy, Delivery, and Early Developmental Milestones
Pregnancies were normal in the births of 80.3% of the children with MPS IIIA (57.1% MPS IIIB). Complications reported during pregnancy were cervix insufficiency (6.2%), infections (5.6%), and vaginal bleedings (4.2%). Previous miscarriages were observed in 29.2% of women (19 of 65). Six patients with MPS IIIA (8.5%) were born prematurely, which is within the range found in the general population.\textsuperscript{24} Normal early developmental milestones were reached by 26.1% of the studied patients with MPS IIIA (14.3% type B). Normal speech development was defined as talking before 15 months and normal motor development as walking independently before 18 months.\textsuperscript{25,26} In 66.7% of the patients with MPS IIIA studied, speech development was delayed (71.4% type B). Two patients with MPS IIIA (7.3 and 8.5 years of age) had never spoken at the time of data collection. Motor development was delayed in 33.3% of the patients with MPS IIIA (50% type B). In 2 patients, parents could remember neither speech nor motor development milestones (excluded from evaluation). A delay of both speech and motor development was observed in 26.1% of the patients with MPS IIIA (35.7% type B). Data on the early development of the patients with MPS IIIA are summarized in Table 3.

First Symptoms and Diagnosis
First symptoms in patients with MPS IIIA were observed, on average, at 7 months of age (SD: 1.2 months [range: 0.0 – 4.5 months]). Although 67.6% of the patients with MPS IIIA first presented symptoms within the first year of life, the average age at diagnosis was 4.5 years (SD: 2.6 years [range: 0.25 –13.8 years]). One third of the patients with MPS IIIA were diagnosed after 5 years of age. The relative distribution of the patients with MPS IIIB diagnosed at different ages was almost identical to the patients with MPS IIIA. Sleep disturbances and behavioral...
abnormalities (hyperactivity, aggressive behavior, and unawareness of dangerous situations) were both reported as being the first noticeable symptoms of the disease in 38% of the patients with MPS IIIA. One fifth of the patients with MPS IIIA first presented with speech delay (Table 4).

**Clinical Presentation**

Sixty-one patients (85.9%) showed coarse facial features. Ten patients at an age between 6.8 and 32.8 years did not show coarse facial features. Macrocephaly was present in 52 patients (73.2%) and microcephaly in 1 patient (1.4%), whereas 18 patients (25.4%) had a normal head circumference. Hearing impairment was reported in 32 patients (45.1%), and hearing aids were used by 21 of these patients. Eye testing was conducted in 39 (55.0%) of the patients with MPS IIIA. Refraction abnormalities were found in 14 patients (35.9%), myopia was apparent in 11 patients (28.2%), and hyperopia was found in 3 patients (7.7%). Strabismus was found in 2 patients (5.1%). No patients presented with corneal clouding. Heart disease, mainly mitral valve followed by aortic valve defects, was detected in 18 patients (25.4%) at the time of data collection. More than one third (38.0%) of the patients with MPS IIIA were never tested for heart disease. Organ enlargement was observed in 64 patients (90.1%). Hernia occurred in 54.9% of the patients with MPS IIIA. Hyperactivity occurred, on average, at 3.3 years of age (SD: 2.6 years [range: 0–17 years]) and declined at 8.8 years (SD: 2.6 years [range: 4.8–16 years]). Persistent nocturnal enuresis was observed in 74.6%, and persistent diurnal, as well as nocturnal enuresis, was reported in 64.8% of the patients with MPS IIIA. In 37 (52.1%) of the patients with MPS IIIA, seizures were reported at the time of data collection. In 68.1% of patients >10 years of age, epilepsy was found (>15 years: 73.9%; >20 years: 81.8%). Epilepsy first presented, on average, at 10.9 years (SD: 4.4 years [range: 3.0–22.75 years]). Of the 71 patients with MPS IIIA studied, 69 (97.2%) showed mental retardation at the time of data collection.

**Scoring of the Clinical Course**

The clinical course of the disease was tracked using the FPSS, which assessed motor function, speech, and cognitive function (Table 1). The median scores for motor function, speech abilities, and cognitive function of all of the patients with MPS IIIA, by age, are shown in Fig 1. The onset of speech regression (score of 2) in patients with MPS IIIA was first observed at 2.8 years of age (SD: 1.9), whereas motor and cognitive functions began to regress at age 4.1 years (SD: 3.6) and 3.0 years (SD: 1.4), respectively. Severe regression (score of 1) in speech, motor, and cognitive function was found at an average age of 5.7 (SD: 2.7), 9.9 (SD: 4.3) and 8.2 (SD: 3.7) years, respectively. Speech was lost (score of 0) at an age of 8.2 (SD: 3.1) years. Loss of motor and cognitive functions was observed at an average age of 12.4 (SD: 5.3) and 13.1 (SD: 4.2) years, respectively. Scoring profiles for speech in patients with MPS IIIB were similar to those of the patients with MPS IIIA. Motor and cognitive functions in the patients with MPS IIIB started declining at ages similar to those of the patients with MPS IIIA; however, the average age of loss of these functions was reported 5.6 and 5.5 years later, respecti-

---

**TABLE 2** Age and Gender Distribution of the Studied MPS III Population

<table>
<thead>
<tr>
<th>MPS III Subtype</th>
<th>Age (N = 89), n (Male/Female)</th>
<th>Total n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0–5 y</td>
<td>&gt;5–10 y</td>
</tr>
<tr>
<td>A</td>
<td>4 (1/3)</td>
<td>19 (9/10)</td>
</tr>
<tr>
<td>B</td>
<td>2 (1/1)</td>
<td>3 (3/0)</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1 (1/0)</td>
</tr>
<tr>
<td>Total</td>
<td>6 (2/4)</td>
<td>23 (13/10)</td>
</tr>
</tbody>
</table>

Data are for the time of data collection.

---

**TABLE 3** Early Developmental Milestones of Patients With MPS IIIA (N = 69)

<table>
<thead>
<tr>
<th>Status</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal early development</td>
<td>18 (26.1)</td>
</tr>
<tr>
<td>Late talking (≥15 mo)</td>
<td>28 (40.6)*</td>
</tr>
<tr>
<td>Late walking (≥18 mo)</td>
<td>5 (7.2)</td>
</tr>
<tr>
<td>Late talking and walking</td>
<td>18 (26.1)</td>
</tr>
</tbody>
</table>

* Two patients never spoke.

---

**TABLE 4** First Symptoms Observed by Parents of Patients With MPS IIIA (N = 71)

<table>
<thead>
<tr>
<th>First Symptom*</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep disturbances</td>
<td>27 (38)</td>
</tr>
<tr>
<td>Behavioral abnormalities</td>
<td>27 (38)</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>22 (31)</td>
</tr>
<tr>
<td>Recurrent infections</td>
<td>16 (23)</td>
</tr>
<tr>
<td>Speech delay</td>
<td>14 (20)</td>
</tr>
<tr>
<td>Hemia</td>
<td>14 (20)</td>
</tr>
<tr>
<td>Motor function delay</td>
<td>13 (18)</td>
</tr>
<tr>
<td>General developmental delay</td>
<td>10 (14)</td>
</tr>
<tr>
<td>Skeletal abnormalities</td>
<td>9 (13)</td>
</tr>
<tr>
<td>Coarse facial features</td>
<td>7 (10)</td>
</tr>
<tr>
<td>Delay in cognitive development</td>
<td>5 (7)</td>
</tr>
</tbody>
</table>

* Multiple answers were permitted.
tively, than in the patients with MPS IIIA. The loss of all 3 of the assessed abilities, as measured by the TDS, was observed at an average age of 12.5 years (range: 8.0–26.5) in patients with MPS IIIA. In a small group of patients with MPS IIIA (9.9%) speech, motor, and cognitive skills were partially preserved (minimum score of 1 in all of the assessed abilities) to a maximum age of 23.8 years, indicating a slower progression of the disease.

Developmental Regression
One third of the patients with MPS IIIA showed developmental regression before 4 years of age and 78.9% before 6 years of age. The first abilities in which regression were noticed were speech (67.6% of patients) and mental/cognitive abilities (42.3% of patients).

Age at Death
Eight of the patients with MPS IIIA (and 2 with subtype B) had died at the time of data collection. The median age at the time of death was 15.2 years (SD: 5.6 years [range: 8.5–25.5 years]).

DISCUSSION
The present study provides a detailed description of the natural course of MPS III. This report summarizes statistically relevant data on the age of onset, symptoms, diagnosis, and progression of the disease in 71 patients with MPS IIIA. Patient numbers for other subtypes were not large enough to produce statistically relevant data. Baehner et al reported on the incidence of MPS III in Germany by the diagnosis of 211 patients with MPS III between 1980 and 1995 composed of subtype A (149, corresponding with 71%), B (49, corresponding with 23%), and C (13, corresponding with 6%). Numbers and the distribution of subtypes were consistent with the data assessed in this study. Furthermore the present study seems to cover ~40% of all of the patients with MPS III in Germany. Studies using a scoring system for other neurodegenerative lysosomal diseases have been conducted for late infantile and juvenile neuronal ceroid-lipofuscinosis. The FPSS is an reliable and valid instrument, which can measure and document the course of the disease in MPS III. It is easy to use and provides important results concerning the clinical phenotypes and progression of disease. The majority of the patients with MPS IIIA (80%) were born after uneventful pregnancies. There were, however, a surprisingly high number of miscarriages observed in mothers of patients with MPS IIIA (19 of 65 women, corresponding with 29%). The general incidence of miscarriage is age dependent, and miscarriages are found in 10% of 20- to 24-year-old and 15% of 30- to 34-year-old women. Because the exact age of the patients’ mothers at the time of miscarriage is unknown, and no further investigations (eg, chromosomal testing) were conducted, it remains unclear whether the increased rate of miscarriages is related to MPS IIIA. At present, there are no reports of an increased risk of miscarriage in pregnancies involving children with lysosomal storage disorders in the literature. The majority of the patients with MPS IIIA (68%) first presented symptoms of the disease, mainly sleep disturbances and behavioral abnormalities, within the first year of life (Table 4). Patients with MPS IIIB presented first symptoms at a similarly early age. Only 10% of the patients with MPS IIIA presented with coarse facial features as a first symptom in contrast to ~30% of the patients with MPS IIIB. Our data showed that three quarters of children with MPS IIIA demonstrated delays in early developmental milestones (Table 3), the most common being a delay in speech development (67%). Speech delay was equally common in patients with MPS IIIB, but motor developmental delay was more common than in patients with MPS IIIA (33% [type A] vs 50% [type B]). Initial diagnosis in patients with MPS IIIA and B was made at an average age of 4.5 years, most probably because of developmental regression first presenting at this age. One third of patients were diagnosed after
5 years of age. We, therefore, suggest that every child with an unspecified developmental delay, behavioral abnormalities, or a delay in speech development should be tested for MPS III. The clinical course of the disease, measured by the TDS, varied substantially between patients in terms of both the age of onset and progression of the disease. Single scoring curves for motor function, speech abilities, and cognitive function revealed that deterioration in motor function, speech, and cognitive function (score of 2) was observed at an average age of between 2.6 and 4.1 years. Loss of speech (score of 0), however, was observed between 4.4 and 4.9 years earlier than loss of motor function and cognitive function. Speech, motor function, and cognitive skills were lost at a mean age of 12.5 years. In a small population of patients >12.5 years (9.9%), however, speech and motor skills were partially preserved (minimum score of 1 in all of the assessed abilities) to a maximum age of 23.8 years. This indicates that the FPSS may be used to classify patients into groups with a rapid or slower course of the disease, which may have an important impact on parental counseling, as well as therapeutic interventions (eg, surgical procedures). One third of the patients with MPS IIIA who were involved in the study were never tested for heart disease. Although the majority of patients did not show significant heart disease with therapeutic consequences, it is important to check for heart valve disease, because it is an indication for endocarditis prophylaxis.

CONCLUSIONS
To our knowledge, this is the largest study conducted in patients with MPS IIIA and the first assessment of the natural course of the disease using a scoring system. The FPSS may be used to predict the possible course of the disease and might also be used as a basis for the evaluation of the effectiveness of future therapies.

ACKNOWLEDGMENTS
The study was supported financially and technically by the Gesellschaft für Mucopolysacharidosen e.V. Germany.

We thank the patients’ families for participating and for making the study possible. We also thank our colleagues for referring patients and Simon Adams for help with the article.

REFERENCES
22. De Jong JG, Wevers RA, Laarakkers C, Poorthuis BJ. Dimeth-


Aerobic Fitness Attenuates the Metabolic Syndrome Score in Normal-Weight, at-Risk-for-Overweight, and Overweight Children

Katrina D. DuBose, PhD\(^a\), Joey C. Eisenmann, PhD\(^b\), Joseph E. Donnelly, EdD\(^c\)

\(^a\)Department of Exercise and Sport Science, East Carolina University, Greenville, North Carolina; \(^b\)Department of Health and Human Performance, Iowa State University, Ames, Iowa; \(^c\)Schiefelbusch Institute for Life Span Studies, Center for Physical Activity and Weight Management, University of Kansas, Lawrence, Kansas

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The purpose of this study was to examine the combined influence of aerobic fitness and BMI on the metabolic syndrome score in children.

METHODS. A total of 375 children (193 girls and 182 boys) aged 7 to 9 years were categorized as being normal weight, at risk for overweight, and overweight on the basis of BMI and aerobic fitness (high or low based on median split) via a submaximal physical working capacity test. Participants were cross-tabulated into 6 BMI fitness categories. High-density lipoprotein cholesterol and triglyceride levels, homeostasis assessment model of insulin resistance, mean arterial pressure, and waist circumference were used to create a continuous metabolic syndrome score.

RESULTS. Both BMI and fitness were associated with the metabolic syndrome score. In general, the metabolic syndrome score increased across the cross-tabulated groups with the normal-weight, high-fit group possessing the lowest metabolic syndrome score and the overweight, unfit group possessing the highest metabolic syndrome score. Children who were at risk for overweight and had high fitness had a lower metabolic syndrome score compared with those at-risk-for-overweight, less-fit children, and the score was similar to that of the less-fit, normal-weight children. Furthermore, a high fitness level resulted in a lower metabolic syndrome score in overweight children compared with overweight children with low fitness.

CONCLUSIONS. High fitness levels modified the impact that BMI had on the metabolic syndrome score in children. Increasing a child’s fitness level could be one method for reducing the risk of obesity-related comorbidities.
The metabolic syndrome consists of a clustering of specific cardiovascular disease risk factors that include elevated blood pressure, dyslipidemia, hyperglycemia or insulin resistance, and central obesity. Depending on the definition used for the metabolic syndrome, ~35% to 39% of adults in the United States have the metabolic syndrome. Given the epidemic of pediatric obesity and emergence of type 2 diabetes among adolescents, it should not be surprising that the metabolic syndrome has also been reported to exist in children and adolescents. Estimated prevalence rates range from 5% to 39% in the general population of youth and overweight adolescents, respectively.

As mentioned above, obesity has reached epidemic proportions in youth. This excess amount of body weight (ie, BMI) is positively related to the metabolic syndrome in adults and youth. Furthermore, a prospective study reported that a pronounced increase in BMI during adolescence was associated with a greater likelihood of possessing the metabolic syndrome in adulthood, which is especially problematic for youth, because the metabolic syndrome and individual components track from adolescence to adulthood.

In adults, high fitness levels are inversely associated with the metabolic syndrome, even among overweight and obese individuals. High fitness levels are also related to a better metabolic syndrome profile among adolescents. Moreover, adolescents who maintain a high fitness level through early adulthood are less likely to develop the metabolic syndrome in young adulthood.

In adults, research has shown that even among those who are obese there is a reduced risk for chronic disease and mortality if they have a high fitness level or participate in regular physical activity. In adolescents and children, studies have shown that BMI is positively related to the metabolic syndrome whereas fitness levels are inversely related to the metabolic syndrome. Until recently, the combined influence of BMI and fitness level on metabolic syndrome components in youth was unknown. Eisenmann et al examined this relationship in adolescence and reported that a high aerobic fitness level and a high BMI resulted in a lower metabolic syndrome score compared with the low-fitness and high-BMI group. However, these studies included mainly normal-weight adolescents.

Therefore, the purpose of this study was to examine the independent and combined influences of BMI and aerobic fitness on the metabolic syndrome in children. Furthermore, we examined whether the metabolic syndrome score varied among normal-weight, at-risk-for-overweight, and overweight children by aerobic fitness level. It was hypothesized that aerobic fitness would attenuate the metabolic syndrome score within BMI categories.

METHODS

Participant Recruitment
The participants used for this study were part of a 3-year physical activity intervention called Physical Activity Across the Curriculum. However, only baseline measures (ie, before random assignment into the intervention) are included in this report. A subsample of second- and third-grade children (aged 7–9 years old) from 22 elementary schools was recruited for additional baseline testing. Inclusion into the subsample consisted of the following criteria: (1) both the parent and child gave written consent and assent, respectively, to participate in the testing in accordance with the human subjects committee at the university; (2) the child had to participate in all of the tests (ie, the child could not choose which tests to complete); and (3) the child did not have insulin-dependent diabetes, cardiovascular disease, or any other disease that limited physical activity participation.

A total of 852 children volunteered for the baseline testing from a possible 2494 children. Because of time constraints, all of the children could not participate, so a random sample of 499 second- and third-graders was selected to participate in the baseline testing. Approximately equal numbers of boys and girls participated, and 27% of the sample was a race other than non-Hispanic white. Baseline testing included the following procedures: height, weight, circumference measurements, skinfold measurements, resting blood pressure, fasting blood draw, aerobic fitness and academic achievement tests, and a physical activity and nutritional surveys. For this analysis, the variables of interest included height, weight, waist circumference, blood pressure, blood chemistry, and aerobic fitness.

Anthropometric Measures
Height was measured to the nearest 0.1 cm using a portable stadiometer (model IP0955, Invicta Plastics Limited, Leicester, England). Weight was measured to the nearest 0.1 kg using a portable electronic scale (model 68987, Belour Inc, Saukville, WI). Both height and weight were measured in duplicate with shoes off but while subjects were wearing lightweight clothing. BMI was calculated as weight in kilograms divided by height in meters squared, and the children were grouped into 1 of 3 BMI categories: (1) normal weight (<85th percentile); (2) at risk for overweight (85th–94th percentile); or (3) overweight (≥95th percentile) according to the age- and gender-specific reference values of the Centers for Disease Control and Prevention.

Waist circumference was measured in duplicate to the nearest 0.1 cm using a Gullick tape at the smallest girth around the trunk in the horizontal plane underneath the participant’s clothing. There was no difference for intertester reliability for waist circumference (P = .55), and the coefficient of variation was 1.65%. Skinfold...
measurements were taken in duplicate on the right side of the body using procedures outlined by the American College of Sports Medicine at the calf and triceps. Percentage of body fatness was estimated from skinfold measurements using the equation by Lohman and Going. The correlation between BMI and percentage of body fat was .80; therefore, BMI categories were used given their clinical and epidemiologic use.

**Blood Pressure**

Resting blood pressure was measured in duplicate by trained personnel using a random-zero sphygmomanometer (Hawksley & Sons Ltd, Lancing, United Kingdom) according to standard methods. To determine the appropriate cuff size, the child’s arm circumference was measured. Children rested quietly for 5 minutes before measurement. The first and fifth Korotkoff sounds were recorded as systolic and diastolic blood pressure, respectively. Mean arterial pressure (MAP) was calculated by using the formula MAP = [(systolic blood pressure − diastolic blood pressure)/3] + diastolic blood pressure. Intertester reliability for systolic or diastolic blood pressure was similar (P = .14 and .11, respectively), and the coefficient of variation for both measures was 5.3%.

**Blood Collection and Storage**

Blood samples were collected using standard venipuncture methods by a trained phlebotomist after an 8-hour fast. Blood samples were processed at the study site by centrifuging the samples, placing the serum in prelabeled vials, storing the samples at −70°C, and shipping the samples to the University of Colorado Health Sciences Center (Denver, CO) for further processing. Blood samples remained stored at −70°C until analyses were conducted.

**Blood Analyses**

Glucose, total cholesterol, and triglyceride concentrations were measured enzymatically using a Cobas Mira Chemistry System (Roche Diagnostic Systems, Indianapolis, IN). High-density lipoprotein cholesterol (HDL-C) concentrations were also measured enzymatically using a Cobas Mira Chemistry System (Diagnostic Chemicals Ltd, Oxford, CT). Insulin levels were measured using a radioimmunoassay (Diagnostic Systems Laboratory, Webster, TX). Homeostasis model assessment (HOMA), an insulin resistance indicator, was calculated as fasting insulin (units per milliliter) × fasting glucose (milligrams per deciliter)/22.5. The coefficient of variation for all of the blood measurements was <5% for both interassay and intraassay quality control.

**Physical Working Capacity**

A modified physical working capacity (PWC) 170-cycle ergometer test assessed aerobic fitness. The PWC has been shown to correlate well with maximal oxygen consumption in boys and girls (r = 0.70 and 0.71, respectively). Before beginning the test, participants were acclimated to the pedaling cadence. During the graded exercise test, participants pedaled on a cycle ergometer until their heart rate reached ≥85% of heart rate reserve or until the participant could no longer maintain a cadence of 60 rpm. The PWC test had a total of four 2-minute stages, and after each stage, the load was increased based on the participant’s heart rate. Heart rate was recorded every minute using a Polar heart rate monitor (Polar Accurex Plus; Polar Electro, Inc, Woodbury, NY). The maximum workload (watts) was recorded and divided by body weight (kilograms) for each participant. Because PWC varies by age and gender, it was regressed onto age and gender. The standardized residuals (z scores) were used to create high and low fitness categories based on the median split.

**Statistical Analysis**

Of the 499 children randomly selected to participate in the baseline testing, 34 did not participate because of absences on the testing day, 10 ate the morning of the blood draw, 1 did not have waist circumference, 76 had incomplete blood data, and 3 did not complete the PWC, leaving 375 children (193 girls and 182 boys; 272 white, 42 Hispanic, 23 black, 9 Asian, 5 Native American, 1 Pacific Islander, 16 >1 race, and 7 unknown or not reported) with complete data for statistical analysis. Sample bias was not present, because demographic characteristics (ie, age, gender, race, and BMI) were similar between those children who were either included or excluded from the data analyses. Currently, there is no universally accepted definition for the metabolic syndrome in children. Therefore, a continuous metabolic syndrome score was created. The metabolic syndrome score was derived by first standardizing the individual metabolic syndrome variables (waist circumference, MAP, HOMA, and HDL-C and triglyceride levels) by regressing them onto age, gender, and race to account for age-, gender- and race-related differences. The standardized HDL-C level was multiplied by −1, because it is inversely related to the metabolic syndrome risk. The standardized residuals (z scores) for the individual variables were summed to create the continuous metabolic syndrome score. These variables were chosen because they represent the same variables (except blood glucose) used in the adult clinical criteria for the metabolic syndrome. HOMA was chosen instead of glucose because most children have normal fasting glucose, and HOMA is related to insulin resistance. MAP was used instead of systolic and diastolic blood pressure to ensure that blood pressure was represented similar to the other factors, and MAP includes both systolic and diastolic pressures in its calculation. To date, it is unknown whether one factor is more important than another for the development of the metabolic syndrome; thus, each.
factor was weighted equally. A higher metabolic syndrome score indicates a less favorable metabolic profile.

Descriptive statistics were calculated for the total sample and by gender. Gender differences in descriptive variables were determined by an independent t test. Pearson’s correlations examined the associations among BMI, PWC, and the metabolic syndrome. An analysis of covariance evaluated independent differences between BMI and PWC categories on the metabolic syndrome score, controlling for age, gender, and race. Because BMI was used as a predictor variable, and BMI and waist circumference are highly correlated ($r = 0.87$ in this sample), the analyses were conducted with and without including waist circumference in the derivation of the score. To test the main hypothesis, 6 BMI-PWC (fat-fit) groups were created: (1) normal weight, high PWC; (2) normal weight, low PWC; (3) at risk for overweight, high PWC; (4) at risk for overweight, low PWC; (5) overweight, high PWC; and (6) overweight, low PWC. Differences in the metabolic syndrome score among the fat-fit groups were examined by analysis of covariance, controlling for age, gender, and race. Posthoc comparisons were made using Tukey least-significant difference. Statistical significance was set at $P < .05$.

**RESULTS**

Descriptive statistics for the total sample and by gender are shown in Table 1. Approximately 21% of the children were at risk for overweight, and an additional 23% were overweight. Girls had higher BMI and body fat percentage and lower PWC values compared with boys ($P < .05$). The metabolic variables were similar between the genders.

Univariate analyses examining the relationships between fatness and fitness (as measured by BMI) and the metabolic syndrome score showed that the correlations were stronger between BMI and the metabolic syndrome score when waist circumference was included in the score ($r = 0.70$) compared with when it was excluded from the score ($r = 0.51$). In contrast, the correlations between PWC and the metabolic syndrome score were similar when waist circumference was included ($r = -0.46$) and excluded ($r = 0.41$) in the metabolic syndrome score. There were significant differences in the metabolic syndrome score between BMI groups (Table 2) and fitness groups (Table 3). There was a graded relationship across the BMI groups when the normal weight group had the lowest metabolic syndrome score and the overweight group had the highest metabolic syndrome score ($P < .05$). The metabolic syndrome score was significantly lower in the low-fitness group compared with the high-fitness group ($P < .05$).

Figure 1 shows the differences in the metabolic syndrome score across the BMI aerobic fitness groups. It is noteworthy to mention the percentages of children in each category were as follows: 36% normal weight, high fitness; 19% normal weight, low fitness; 8% at risk for overweight, high fitness; 13% at risk for overweight, low fitness; 5% overweight, high fitness; and 18% overweight, low fitness. Furthermore, 35% of children in the normal-weight group possessed low fitness, whereas 30% of the at-risk-for-overweight and overweight children possessed high fitness. In general, the metabolic syndrome score increased across groups, with the normal-weight, high-fitness group possessing the lowest metabolic syndrome score and the overweight, unfit group possessing the highest metabolic syndrome score. Several significant differences existed between the 6 fat-fit groups. Of particular note are the differences within BMI groups by fitness level and the comparison of values between the children in the normal-weight, low-fit group and the at-risk-for-overweight children with high fitness. The metabolic syndrome score was significantly lower in children in the high-fitness group in all 3 of the BMI groups compared with their low-fitness counterparts within the same BMI group ($P < .05$). Furthermore, the metabolic syndrome score was not significantly different between the normal-weight, low-fitness

### Table 1: Descriptive Statistics of the Sample According to Gender

<table>
<thead>
<tr>
<th>Variable</th>
<th>Boys ($n = 183$), Mean (SE)</th>
<th>Girls ($n = 192$), Mean (SE)</th>
<th>Total ($n = 375$), Mean (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>7.7 (0.7)</td>
<td>7.7 (0.7)</td>
<td>7.7 (0.7)</td>
</tr>
<tr>
<td>Height, cm</td>
<td>130.4 (6.1)</td>
<td>130.1 (7.0)</td>
<td>130.2 (6.6)</td>
</tr>
<tr>
<td>Mass, kg</td>
<td>30.1 (7.0)</td>
<td>31.2 (9.0)</td>
<td>30.6 (8.1)</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>17.5 (3.0)a</td>
<td>18.2 (3.7)</td>
<td>17.9 (3.4)</td>
</tr>
<tr>
<td>WC, cm</td>
<td>59.5 (8.0)</td>
<td>59.3 (8.7)</td>
<td>59.4 (8.4)</td>
</tr>
<tr>
<td>Body fat %</td>
<td>17.7 (8.3)a</td>
<td>22.2 (7.0)</td>
<td>20.0 (8.0)</td>
</tr>
<tr>
<td>SBP, mm Hg</td>
<td>106.5 (9.4)</td>
<td>104.7 (10.0)</td>
<td>105.6 (9.7)</td>
</tr>
<tr>
<td>DBP, mm Hg</td>
<td>67.4 (8.0)</td>
<td>67.3 (8.0)</td>
<td>67.3 (8.2)</td>
</tr>
<tr>
<td>MAP, mm Hg</td>
<td>80.4 (7.2)</td>
<td>79.7 (7.8)</td>
<td>80.1 (7.5)</td>
</tr>
<tr>
<td>HOMA</td>
<td>1.32 (1.64)</td>
<td>1.58 (2.06)</td>
<td>1.45 (1.87)</td>
</tr>
<tr>
<td>TC, mg/dL</td>
<td>179.6 (27.7)</td>
<td>183.4 (33.1)</td>
<td>181.5 (30.6)</td>
</tr>
<tr>
<td>HDL-C, mg/dL</td>
<td>56.6 (11.6)</td>
<td>54.4 (11.0)</td>
<td>55.5 (11.3)</td>
</tr>
<tr>
<td>LDL-C, mg/dL</td>
<td>108.9 (22.7)</td>
<td>113.9 (29.3)</td>
<td>111.4 (26.4)</td>
</tr>
<tr>
<td>VLDL-C, mg/dL</td>
<td>14.1 (5.9)</td>
<td>15.1 (5.4)</td>
<td>14.6 (5.6)</td>
</tr>
<tr>
<td>TGs, mg/dL</td>
<td>70.4 (29.3)</td>
<td>75.5 (26.9)</td>
<td>73.0 (28.2)</td>
</tr>
<tr>
<td>PWC, W/kg</td>
<td>2.78 (0.81)a</td>
<td>2.48 (0.73)</td>
<td>2.63 (0.78)</td>
</tr>
</tbody>
</table>

**WC** indicates waist circumference; **SBP**, systolic blood pressure; **DBP**, diastolic blood pressure; **TC**, total cholesterol; **LDL-C**, low-density lipoprotein cholesterol; **VLDL-C**, very-low-density lipoprotein cholesterol; **TG**, triglyceride.

### Table 2: Metabolic Syndrome Score According to BMI Category

<table>
<thead>
<tr>
<th>Variable</th>
<th>Normal Weight ($n = 212$), Adjusted Mean (SE)</th>
<th>At Risk for Overweight ($n = 79$), Adjusted Mean (SE)</th>
<th>Overweight ($n = 87$), Adjusted Mean (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MetSyn score</td>
<td>$-1.42 (1.8)a,b$</td>
<td>$0.15 (2.0)c$</td>
<td>$3.33 (3.4)$</td>
</tr>
<tr>
<td>MetSyn score without WC</td>
<td>$-0.88 (1.7)a,b$</td>
<td>$0.04 (2.0)c$</td>
<td>$2.11 (2.7)$</td>
</tr>
</tbody>
</table>

WC indicates waist circumference; MetSyn, metabolic syndrome.

* Normal weight is significantly different from at risk for overweight ($P < .05$).
* Normal weight is significantly different from overweight ($P < .05$).
* At risk for overweight is significantly different from overweight ($P < .05$).
group and the at-risk-for-overweight, high-fitness group or the at-risk-for-overweight, low fitness group and the overweight, high-fitness group.

**DISCUSSION**

The purpose of this study was to examine the independent and combined influences of aerobic fitness and fatness (measured by BMI) on the metabolic syndrome in children. Fatness was positively related to the metabolic syndrome score, indicating that excess weight is associated with the metabolic syndrome score. In addition, high fitness was inversely associated with the metabolic syndrome score. Moreover, the metabolic syndrome score varied by BMI fitness groups, where high fitness attenuated the metabolic syndrome score within BMI categories.

Previous studies have reported a relatively high prevalence of the metabolic syndrome in overweight adolescents (eg, 35%–40%).4,6 In the present study, the metabolic syndrome was positively related to BMI and clinical BMI groups (normal weight, at risk for overweight, and overweight), and this relationship was observed with and without the inclusion of waist circum-
ference in the risk score. Fatness and the metabolic syndrome score have been shown to track from early childhood to adulthood12,32; therefore, interventions should target children who have an increased obesity risk.

Aerobic fitness is inversely related to the metabolic syndrome in adults.13,14,33 Furthermore, there is mounting evidence for a similar relationship between fitness and the metabolic syndrome in adolescents15,16,21; however, limited research exists in children <10 years old. The results from the present study showed that fitness was inversely associated with the metabolic syndrome score. Results from the European Youth Heart Study have also shown a similar inverse relationship between aerobic fitness and the clustering of cardiovascular disease risk factors in children.16,34 These results indicate that promoting the development of aerobic fitness during childhood is important to reduce the risk of developing adverse cardiovascular disease risk factors.

Although a growing body of evidence is being established for the independent associations between fitness and fatness on the metabolic syndrome in children, the combined influence of fitness and fatness on the metabolic syndrome has received limited attention. The present study showed that the metabolic syndrome score varied by clinical cut points of BMI and fitness category. More specifically, high fitness attenuated the metabolic syndrome score within the BMI categories, and this difference was most pronounced in the overweight group. These results are similar to previous studies among adolescents.21,22 Data from the Québec Family Study showed that adolescents with low BMI and high fitness had the lowest metabolic syndrome score, whereas those with high BMI and low fitness had the highest metabolic syndrome score.21 Eisenmann et al22 studied the combined influence of aerobic fitness and percentage of body fat on cardiovascular disease risk factors in Australian youth (9- to 15-year-olds). They found that the high-fat and low-fitness group had a higher metabolic syndrome score compared with the low-fat and high-fitness group. Previous studies examined the relationship among fitness, fatness, and the metabolic syndrome in children and adolescents together; thus, it was unclear whether the relationship would exist in prepubescent children only. Considering the available evidence, it is becoming clear that when children and adolescents are cross-tabulated into fat-fit categories, fitness attenuates the metabolic syndrome score among overweight children and adolescents. Although the reasons for these observations have not been fully explored, they possibly involve genetics, adipocytokines, and oxidative capacity of skeletal muscle.35–37

A finding of particular interest in the present study was the magnitude of the difference in the metabolic syndrome score between the high-fit and low-fit overweight children. Furthermore, the lack of statistical dif-
ference between the normal-weight, low-fit and at-risk-of-overweight, high-fit groups also is important. These findings, along with the attenuation of the metabolic syndrome score within obesity levels as a result of being aerobically fit, further support the idea of being “fat but fit” and having a reduced chronic disease risk compared with those who are fat but unfit.

**CONCLUSIONS**

The results of this study show that the metabolic syndrome is a complex phenotype associated with both fatness and fitness. In addition, the metabolic syndrome score varies within clinical BMI categories (ie, normal weight, at risk for overweight, and overweight) by fitness level. Therefore, aerobic fitness level should be considered when interpreting the metabolic syndrome profile in children. For example, a child who is overweight but participates in activities (eg, swimming, soccer, etc) that positively impact physical fitness may have fewer risk factors compared with a child who is overweight but does not engage in activities that improve physical fitness. This information would be helpful to health care providers when determining treatment options. Aerobic fitness should also be promoted to reduce the risk of obesity-related comorbidities. Longitudinal studies of the fat-fit phenotype during childhood and adolescence and into adulthood are necessary to examine the risk of developing the metabolic syndrome, atherosclerosis, and type-2 diabetes.

**ACKNOWLEDGMENT**

This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases (National Institutes of Health, Bethesda, MD) grant NIH20 DK 061489.

**REFERENCES**


23. Eisenmann JC, Welk GJ, Wickel EE, Blair SN. Combined influence of cardiorespiratory fitness and body mass index on cardiovascular disease risk factors among 8–18 year old youth:
the Aerobics Center Longitudinal Study. Int J Pediatr Obes. 2007;2:66–72
Is Childhood Vaccination Associated With Asthma? A Meta-analysis of Observational Studies

Ran D. Balicer, MD, MPHa, Itamar Grotto, MD, MPHb, Marc Mimouni, MDb,c, Daniel Mimouni, MDc,d

aDepartment of Epidemiology, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel; bSchneider Children’s Medical Centre of Israel, Petach Tikva, Israel; cSackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; dDepartment of Dermatology, Rabin Medical Center, Petach Tikva, Israel

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

BACKGROUND. The possible link between immunization and atopic diseases has been under intense debate in the last decade.

OBJECTIVE. The aim of this study was to systematically review the available evidence on the association of whole-cell pertussis and BCG vaccination with the risk of asthma in childhood and adolescence.

METHODS. The major medical electronic databases (Medline, National Library of Medicine Gateway, and Cochrane Library) were searched, and reference lists of the relevant publications were reviewed for relevant birth-cohort studies and randomized, controlled trials from 1966 to March 2006. Only studies that directly compared vaccinated and unvaccinated children, validated vaccination status by medical charts, and used preset criteria to define asthma were included. Data were abstracted by using a standardized protocol and computerized report form. Results were analyzed by applying a fixed-effect or random-effect model, according to the heterogeneity of the studies. Sensitivity analyses by scoring criteria were performed.

RESULTS. Seven studies of pertussis vaccination (with a total of 186,663 patients) and 5 studies of BCG vaccination (with a total of 41,479 patients) met our inclusion criteria. No statistically significant association was detected between either whole-cell pertussis or BCG vaccination and incidence rates of asthma during childhood and adolescence. This lack of a significant association proved to be robust on sensitivity analyses for BCG but not for pertussis vaccine.

CONCLUSIONS. Currently available data, based on observational studies, do not support an association, provocative or protective, between receipt of the BCG or whole-cell pertussis vaccine and risk of asthma in childhood and adolescence.
In the past 3 decades we have witnessed a dramatic increase in the prevalence of asthma and allergic disease worldwide, most notably in countries with a Western lifestyle. Most cases of asthma first appear in childhood, with 80% to 90% of patients diagnosed by 6 years of age. Gene-environment interactions are critical to the pathogenesis of allergic disorders such as asthma, and the susceptibility to asthma was shown to be increased by conditions that are present early in life, including family history of asthma or other manifestations of atopy, male gender, low birth weight, preterm birth, young maternal age, early cessation of breastfeeding, and parental smoking.

The association between vaccination uptake and the risk of atopic diseases was first proposed by Odent et al in a letter published in the Journal of the American Medical Association in 1994. Thereafter, several studies have suggested either a provocative or protective effect of immunization depending on the specific vaccine, the target population, and the age at which the vaccine was administered. Whole-cell pertussis (as a component of diphtheria-tetanus-pertussis [DTP] vaccines) and bacille Calmette-Guérin (BCG) vaccines are among the most extensively studied in regard to their suggested effect on risk of atopic diseases.

More-recent systematic reviews have failed to support these findings (although no formal meta-analyses have been conducted, partly because of the heterogeneity of the studies), including 2 independent population-based studies published in 2005 that reported a possible atopy-protective effect of immunization. It is important that researchers clarify this issue, because unless refuted, the perception that immunization causes asthma may become a significant determinant of parents' attitudes toward routine vaccination of their children.

The aim of this study was to systematically review the literature on the possible association of whole-cell pertussis and BCG vaccination in the first year of life with the incidence of asthma in childhood and adolescence and to perform a meta-analysis of the relevant studies.

**MATERIALS AND METHODS**

**Search Strategy**

A search of the major medical electronic databases (National Library of Medicine Gateway, Medline, and Cochrane Library) was conducted for articles published from January 1966 to March 2006 that contained the keywords “vaccine,” “BCG,” “pertussis,” “allergy,” “atopy,” “wheezing,” and “asthma,” alone or in combination. Thesaurus and free-text terms (including synonyms) were used in various combinations depending on the requirements of the particular database. In addition, we used the “related articles” function in PubMed, as well as the “articles citing this study” function in online journals. We also manually searched the reference lists in all identified publications and recent systematic reviews.

**Selection of Articles**

The abstracts of all articles identified were reviewed, and 71 relevant original scientific studies were read in full. In the preliminary assessment, we excluded cross-sectional studies and studies that did not directly compare subjects who were vaccinated with BCG or DTP/whole-cell pertussis vaccine (“pertussis vaccine”) with unvaccinated subjects. We also excluded studies in which asthma status in childhood or early adolescence (up to 16 years old) did not serve as an independent outcome measure. When several studies pertained to the same cohort, the study with the longest follow-up period was selected for inclusion. The 24 remaining articles for final assessment were independently reviewed and scored by 2 of us (Drs Balicer and Grotto).

**Data Abstraction and Validity Assessment**

Two of us (Drs Balicer and Grotto) abstracted information from each of the 24 selected studies. All data were abstracted by using a standardized protocol and computerized report form.

Among these 24 articles, we selected only those that met our preset criteria: randomized, controlled trial or birth-cohort study, either prospective or retrospective (including those with a nested case-control design); assessment of vaccination status using the medical charts; asthma/bronchial hyperresponsiveness diagnosed on the basis of a validated questionnaire or the medical charts; and asthma defined as (1) at least 1 reported or recorded episode of wheezing or bronchial obstruction, (2) reported coughing in the morning or during the day or evening in the autumn and winter and coughing daily for ~3 months/year, (3) physician-diagnosed asthma or obstructive bronchitis, or (4) ever-recorded medications for asthma or wheezing. Thirteen articles were eliminated on this basis, which left 11 studies for analysis.

**Quantitative Data Synthesis**

All analyses were performed separately for pertussis and BCG vaccination. The overall odds ratios (ORs) and 95% confidence intervals (CIs) for asthma were calculated by using the OR and variance of each study. The overall measure was summarized by the precision-based estimates described by Fleiss and Kleinbaum et al, which assume a homogeneity of effect between studies (fixed-effect model), and the method of DerSimonian and Laird, which factors in both within-study variance and heterogeneity between studies (random-effect model). Heterogeneity of the ORs across n studies was tested with the following formula: $\chi^2$ heterogeneity = $\sum w_i M_i^2 - (\sum w_i M_i)^2 / \sum w_i$, where $M_i$ is an individual measure of association (logarithm of the OR), and $w_i$ is a weighting.
factor equal to the reciprocal of the squared SE (determined by the upper and lower limits of the 95% CI) of the individual measure. Statistical significance was evaluated with $n - 1$ degrees of freedom.

Quality Ranking and Sensitivity Analysis

The studies included in the final assessment were appraised and ranked for methodologic quality within the process of the sensitivity analyses. Criteria concerning the study design merited 1 point for each of the following: group randomization, retrieval of outcome data from medical charts, a relatively large unvaccinated group (>25% of the study population), and studies that performed and reported the results of a multivariate analysis that accounted for gender, family history of asthma and/or allergy, socioeconomic status, parental smoking, low birth weight, young maternal age, and prolonged breastfeeding (1 point for each factor adjusted for). For cases in which a relevant multivariate analysis was not performed or its results were not available, univariate analysis results were used in the meta-analysis and no quality score was granted for adjustment criteria. Studies were scored and ranked according to the sum of points (range: 0 – 11). In case of equal scores, the study that scored higher in the first 3 above-mentioned study-design criteria was ranked higher. Disagreements or uncertainties were resolved by discussion among all the investigators.

We performed sensitivity analyses in which we excluded studies from the pool to examine their effect on the pooled estimate and CIs. First, we removed each of the studies, and then we excluded the study with the lowest-quality score and then 2 studies with the lowest-quality scores. We also recalculated the pooled estimates, including only studies in which unvaccinated subjects comprised at least 25% of the total cohort, thereby adjusting for several potential bias factors. All computations were performed by using PEPI 3 (USD, Inc, Stone Mountain, GA) for epidemiologic analysis.

### RESULTS

#### Trial Flow

After screening >2000 potentially relevant citations, 71 reports of relevant original data were identified and read completely, of which 24 met the preliminary inclusion criteria, as detailed above. Thirteen of these studies were ultimately excluded for the reasons detailed in Table 1.

#### Study Characteristics

Seven studies of the pertussis vaccine (with a total of 186,663 patients) and 5 studies of the BCG vaccine (with a total of 41,479 patients) were included in the meta-analysis.13–23

One study assessed both pertussis and BCG vaccines and was included in both analyses.19 The main features of these studies, organized by quality score, are shown in Table 2.

#### Quantitative Data Synthesis and Sensitivity Analyses

**BCG Vaccination and Asthma**

Figure 1 presents the ORs and 95% CIs of the 5 studies that investigated the association between BCG vaccination and asthma. Two of them,19–21 including a relatively large-scale study that yielded statistical results of borderline significance, reported a protective association.19 The studies were not found to be heterogeneous ($\chi^2 = 4.89; P = .299$); therefore, we used the OR that was calculated according to the fixed-effect model as the overall estimate. The OR was 0.98 (95% CI: 0.88 – 1.08), which indicates that BCG vaccination had no overall effect on the occurrence of asthma in childhood or adolescence.

Because the heterogeneity did not reach statistical significance after exclusion of each of the studies, we continued to use the fixed-effect model for our sensitivity analyses (Table 3). The exclusion of each study yielded similar results, with an OR of 0.89 to 1.06 and a CI that overlapped 1.0 in all analyses. Table 3 summarizes these results, as well as 2 additional sensitivity

<table>
<thead>
<tr>
<th>Trial</th>
<th>Year</th>
<th>Country</th>
<th>Vaccine</th>
<th>Main Reason for Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odent et al27</td>
<td>1994</td>
<td>United Kingdom</td>
<td>Pertussis</td>
<td>No validation of vaccination by medical charts</td>
</tr>
<tr>
<td>Kemp et al24</td>
<td>1997</td>
<td>New Zealand</td>
<td>Pertussis</td>
<td>Missing disease-specific data on asthma</td>
</tr>
<tr>
<td>Strannegård et al25</td>
<td>1998</td>
<td>Sweden</td>
<td>BCG</td>
<td>No validation of vaccination by medical charts</td>
</tr>
<tr>
<td>Hurwitz and Morgenstern26</td>
<td>2000</td>
<td>US</td>
<td>Pertussis</td>
<td>No validation of vaccination by medical charts</td>
</tr>
<tr>
<td>Pahari et al22</td>
<td>2002</td>
<td>United Kingdom</td>
<td>BCG</td>
<td>No validation of vaccination by medical charts</td>
</tr>
<tr>
<td>McKeever et al24</td>
<td>2002</td>
<td>United Kingdom</td>
<td>pertussis</td>
<td>Survival-analysis design</td>
</tr>
<tr>
<td>Bager et al23</td>
<td>2003</td>
<td>Denmark</td>
<td>BCG</td>
<td>Older age groups, asthma-definition standard not met</td>
</tr>
<tr>
<td>Benke et al28</td>
<td>2004</td>
<td>Australia</td>
<td>Pertussis, BCG</td>
<td>No validation of vaccination by medical charts</td>
</tr>
<tr>
<td>Da Cunha et al25</td>
<td>2004</td>
<td>United Kingdom</td>
<td>BCG</td>
<td>No validation of vaccination by medical charts</td>
</tr>
<tr>
<td>Martignoni et al22</td>
<td>2005</td>
<td>France</td>
<td>Pertussis, BCG</td>
<td>No vaccine-specific data, no asthma-specific data</td>
</tr>
<tr>
<td>Enríquez et al22</td>
<td>2005</td>
<td>US</td>
<td>Pertussis, BCG</td>
<td>No validation of vaccination by medical charts</td>
</tr>
<tr>
<td>García-Marcos29</td>
<td>2005</td>
<td>Spain</td>
<td>BCG</td>
<td>No validation of vaccination by medical charts</td>
</tr>
<tr>
<td>Bernsen et al33</td>
<td>2006</td>
<td>Netherlands</td>
<td>pertussis</td>
<td>No validation of vaccination by medical charts</td>
</tr>
</tbody>
</table>
analyses including only the 3 or 4 studies with the highest methodologic quality. No difference was noted in either the heterogeneity between studies or the overall measure.

**Pertussis Vaccination and Asthma**

Figure 2 presents the ORs and 95% CIs of the 7 studies on pertussis vaccination and asthma. Two studies reported that vaccination induced asthma,14,17 but in only 1 of them did the between-group difference reach statistical significance.14 The heterogeneity of the 7 studies was found to be statistically significant (\(\chi^2 = 13.98; P = .03\)); therefore, the OR was calculated according to the random-effect model to determine the overall estimate. The OR was 0.99 (95% CI: 0.78–1.25), which indicates that pertussis vaccination had no detectable adverse effect on asthma. Additional calculations using the fixed-effect model yielded similar results (OR: 1.02 [95% CI: 0.92–1.13]).

Because the heterogeneity consistently reached statistical significance with exclusion of each of the studies, we continued to use the random-effect model for our sensitivity analysis. The results remained similar after exclusion of each study, with an OR of 0.93 to 1.1 and a CI that overlapped 1.0 in all analyses. Table 3 summarizes these results, as well as an additional 2 sensitivity analyses including only the 6 studies with the highest-quality scores. There was no change in heterogeneity or in the overall measure. Including only the 5 studies with the highest-quality scores reduced the heterogeneity below the level of significance (\(\chi^2 = 6.62; P = .157\)). In this case, when we used the fixed-effect model, pertussis vaccination had a borderline significant provocative effect on asthma in childhood (OR: 1.26 [95% CI: 1.04–1.54]).

**DISCUSSION**

This meta-analysis was conducted to clarify the controversial findings for the association of immunization and...
risk of atopy. Some vaccines were implicated as having an inciting effect (most notably whole-cell pertussis vaccine but also measles-mumps-rubella vaccine), whereas others were suggested to have a protective effect (most notably BCG vaccine). To cope with the suboptimal methodologies and the multiple outcomes, we selected a single outcome (asthma), excluded the studies with a cross-sectional design, and included only studies that adhered to standard methods of validation of exposure and predetermined outcome criteria. We also focused only on the 2 most-studied vaccines in this context. Although neither vaccine is used as widely today as it was a decade ago, the BCG vaccine was routinely administered to all children in the United Kingdom and Finland until recently and is still universally used in France, and the DTP (rather than the diphtheria-tetanus-acellular pertussis [DTaP]) vaccines are still widely used outside Europe and North America.

Our final selection comprised 11 studies that included a total of 227,570 subjects (1 of which assessed both vaccines). The results did not support either a protective or provocative effect of BCG or whole-cell pertussis vaccination in the first year of life on the likelihood of acquiring asthma in childhood and adolescence.

In the past 3 decades we have witnessed a spectacular increase in the prevalence of asthma and allergic disease worldwide: >130 million people suffer from asthma, and the numbers are increasing. This quadruple increase is most notable in countries with a Western lifestyle. A critical role for environmental factors in driving

<table>
<thead>
<tr>
<th>Studies (Quality Rank)</th>
<th>Precision-Based Estimated OR (95% CI)</th>
<th>DerSimonian and Laird-Based OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>Degrees of Freedom</td>
</tr>
<tr>
<td>BCG vaccination and asthma</td>
<td>All studies (1–5)$^a$</td>
<td>4.89</td>
</tr>
<tr>
<td></td>
<td>Highest quality (1–4)$^a$</td>
<td>4.49</td>
</tr>
<tr>
<td></td>
<td>Highest quality (1–3)$^a$</td>
<td>0.85</td>
</tr>
<tr>
<td>Pertussis vaccination and asthma</td>
<td>All studies (1–7)</td>
<td>13.98</td>
</tr>
<tr>
<td></td>
<td>Highest quality (1–6)</td>
<td>11.76</td>
</tr>
<tr>
<td></td>
<td>Highest quality (1–5)$^a$</td>
<td>6.62</td>
</tr>
<tr>
<td></td>
<td>Small unvaccinated cohorts excluded (1–3 and 7)</td>
<td>9.07</td>
</tr>
</tbody>
</table>

$^a$ The DerSimonian and Laird test was not performed because statistically significant heterogeneity was not observed.
the expression of asthma and other allergic diseases is considered almost certain. Exposure to high amounts of allergens, such as those derived from house dust mites, cannot explain the large intercountry differences in asthma rates or the rising trends. Concerns have been raised that factors that once provided protection from allergic disease may have been lost from the environment.

In 1989, Strachan7 proposed the hygiene hypothesis, which claims that the apparent rise in the prevalence of allergic disease “could be explained if allergic diseases were prevented by infection in early childhood, transmitted by unhygienic contact with older siblings, or acquired prenatally.” This hypothesis suggested that modern-day cleanliness has decreased microorganism stimulation of the immune system in infants, leading to the persistence of an immature immune response and, consequently, an imbalance in T-helper 1 (Th1) and Th2 immunity, which may lead to atopy. Persistence of the immature immune response may depend on the presence of certain predisposing genes and specific environmental exposures.

More recently, extensive epidemiologic research has shown that exposure to microorganisms or their products may play a role in allergic disease. Indirect support was provided by findings of a considerably lower prevalence of allergic diseases in developing countries and in rural areas within 1 country.41 The overrepresentation of allergic sensitization in firstborn children and the lower rates in children from large families or attending day care also point to a possible protective effect of a frequent exchange of infections.40 Accordingly, exposure to microbial substances in stables and in unpasteurized (raw) milk was found to be inversely related to the development of atopy, and the endotoxin load in the mattresses of children brought up in a rural environment was inversely related to the occurrence of hay fever and grass sensitization.42

The hygiene hypothesis implies that any interventional factor that reduces childhood infections may potentially be associated with an increased incidence of allergic diseases. Several studies have demonstrated a statistically significant14,24,26 or nonsignificant22 tendency toward a higher risk of atopy in individuals who have been immunized with the whole-cell pertussis vaccine, administered in most cases as a component of the DTP vaccine. Of these 3 components, pertussis toxin was shown to cause an increased production of pertussis-specific immunoglobulin E antibodies in both animal and human models. However, in most cases, this was a transient change that was unrelated to the development of atopic conditions.19 In this analysis we were not able
to differentiate the individual effect of the diphtheria and tetanus components on the measured outcomes.

Our finding from this analysis that pertussis vaccination is not related to asthma incidence is in keeping with the only randomized, controlled trial to date that addressed this issue. Nilsson et al.13 randomly assigned 667 Swedish children to 4 groups to receive 2- or 5-component acellular pertussis vaccine, whole-cell pertussis vaccine, or DPT vaccine. The cumulative incidence of asthma at 7 years was similar in all groups after adjustment for family history of atopic disease, environmental smoking at home, and other living conditions. Similar findings were noted in other nonrandomized well-controlled studies that compared pertussis-vaccinated and unvaccinated children. One large-scale ecological study even reported a negative correlation of pertussis immunization and wheezing at the population level.41

The selection process and sensitivity analyses used in this study were designed to address several of the methodologic challenges associated with nonrandomized observational studies of vaccinated and unvaccinated populations. To reduce the effect of recall bias, we excluded 9 studies from researchers who failed to validate vaccination status with medical charts. In addition, the unique attributes of specific unvaccinated subpopulations may also act as confounders or introduce misclassification. For example, parents who elect not to vaccinate their child may practice a naturalistic lifestyle, thus introducing various confounding factors. They may also be less likely to seek medical assistance in events of wheezing/asthma and, therefore, will be underrepresented in studies that use only medical charts to assess outcome. McKeever et al.28 found that nonvaccinated children visited their general practitioner less often than vaccinated children in the same cohort, and the association between vaccination and asthma varied considerably between the subgroups according to their visit frequency. Therefore, we performed a separate analysis of studies in which the unvaccinated children comprised >25% of the study cohort, because nonvaccination in these cases may not have been limited to secluded individuals. Alternatively, the differences in beliefs and lifestyle of this subgroup and the vaccinated population may not have been as profound as expected. We found that the results of this subgroup analysis were in keeping with the others in our study.

The only borderline significant positive association of vaccination with asthma incidence was noted in the analysis of the 5 studies of the highest methodologic quality score. The results showed that pertussis vaccine had a provocative effect on the occurrence of asthma (OR: 1.26 [95% CI: 1.04–1.54]). This finding highlights the importance of performing additional, adequately designed evidence-based studies.

BCG immunization during infancy has been suggested to protect against atopic diseases and asthma. Unlike the pertussis vaccine, BCG immunization involves the inoculation of live mycobacteria. Therefore, it may have a direct effect on the immune system, similar to that of tuberculosis and other bacterial infections, and provide protection against atopy in accordance with the hygiene theory. In animal models, BCG vaccination has resulted in a preferential proliferation of Th1 cells and inhibited subsequent immunoglobulin E antibody formation and allergen-induced airway inflammation, even in the presence of established allergies.44-48 This protective effect was also suggested in several human studies, although in most cases, it did not reach statistical significance. One exception is the large-scale study of Grüber et al.19 who compared German and Dutch children and demonstrated a moderate (and borderline significant) protective effect of BCG vaccination on childhood asthma.

The results of our meta-analysis do not support these earlier findings. Similar to pertussis, we found no statistically significant association between BCG vaccination and subsequent asthma/wheezing in childhood and adolescence.

It is noteworthy that much of the preliminary evidence of such a protective impact was reported in studies that used the response to purified protein derivative as a surrogate marker for successful BCG inoculation.49,50 We did not include them in our analysis, because it is unclear if the decreased tuberculin responses were induced by atopy or if they contributed to the atopy-related Th2-type immune response. In another study in Spanish schoolchildren, which suggested a weak but significant protective effect of BCG immunization against asthma and hay fever, the BCG vaccination status was not validated against the medical charts.51 Nevertheless, we included it in our analysis because the authors, who were aware of the problem, claimed that the effect of this potential misclassification was nondifferential, so that the statistically significant increase in the calculated risk was probably an underestimate. In view of this rationale, we also performed an analysis including this study, and the overall protective effect of BCG vaccination neared but did not reach statistical significance.

Thus, the currently available evidence does not indicate a protective effect of BCG vaccination on asthma incidence, although a tendency for such an effect can be detected and may be affirmed in future randomized, controlled trials.

CONCLUSIONS
The currently available data do not support an association, either provocative or protective, of BCG or whole-cell pertussis vaccination in infancy and risk of asthma in childhood and adolescence. These findings could be used to relieve parental concerns that could otherwise lead to vaccination refusal.

The lack of robustness of the results of the sensitivity
analyses of pertussis vaccination and the multitude of potential biases in studies that have used a birth-cohort design stress the need for additional adequately controlled, large-scale studies.

REFERENCES

2. O’Connell EJ. The burden of atopy and asthma in children. Allergy. 2004;59(suppl 78):7–11
4. Odent MR, Culpin EE, Kimmel T. Pertussis vaccination and asthma: is there a link? JAMA. 1994;272:592–593
31. da Cunha SS, Cruz AA, Dourado I, Barroto ML, Ferreira LD, Rodrigues LC. Lower prevalence of reported asthma in adolescents with symptoms of rhinitis that received neonatal BCG. Allergy. 2004;59:857–862
Evaluation and Treatment of Community-Acquired *Staphylococcus aureus* Infections in Term and Late-Preterm Previously Healthy Neonates

Regine M. Fortunov, MDa,b,c, Kristina G. Hulten, PhDa,d,e, Wendy A. Hammerman, RNa,d,e, Edward O. Mason, Jr, PhDa,d,e, Sheldon L. Kaplan, MDa,d,e

aDepartment of Pediatrics, bSections of Neonatology and dInfectious Diseases, Baylor College of Medicine, Houston, Texas; cDepartment of Neonatology and eInfectious Disease Service, Texas Children’s Hospital, Houston, Texas

Financial Disclosure: Dr Kaplan received a grant from Pfizer for the *Staphylococcus aureus* study. The other authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. We describe the evaluation and treatment of neonatal community-acquired *Staphylococcus aureus* disease in the era of community-acquired methicillin-resistant *S. aureus*.

METHODS. We retrospectively reviewed the evaluation and treatment of 126 community-acquired *S. aureus* infections of term and late-preterm previously healthy neonates who were ≤30 days of age between August 2001 and July 2006 at Texas Children’s Hospital.

RESULTS. *S. aureus* infections included 43 pustulosis, 68 cellulitis/abscess, and 15 invasive infections. We found 84 methicillin-resistant and 42 methicillin-susceptible *S. aureus* isolates. Twenty-one patients received outpatient antibiotics before hospital presentation. Systemic infection evaluation included urine, blood, and cerebrospinal fluid cultures in 79, 102, and 84 neonates, respectively. Culture revealed *S. aureus* urinary tract infections in 1, *S. aureus* bacteremias in 6, and aseptic cerebrospinal fluid pleocytosis of unclear cause in 11 neonates. Physicians admitted 106, transferred 5 to other hospitals, and discharged 15 afebrile patients with topical or oral antibiotics. Clindamycin was the predominant antistaphylococcal intravenous and oral antibiotic for pustulosis and cellulitis/abscess infections. One patient with systemic *S. aureus* and herpes simplex virus infection died. At discharge after inpatient treatment, physicians prescribed no antibiotics for 43 patients and oral or topical antibiotics for 62 patients. Outpatient treatment failed for 1 patient who was discharged after intravenous therapy and was readmitted. Eighty percent (16 of 20) of patients with mastitis alone completed treatment with outpatient oral antibiotics.

CONCLUSIONS. Evaluation and treatment strategies for neonatal community-acquired *S. aureus* disease are varied at our hospital. Prospective studies are needed to determine optimal management strategies.
Staphylococcus aureus is a frequent neonatal pathogen that causes a variety of infections. Community-acquired (CA) methicillin-resistant S aureus (MRSA) isolates have been documented in nosocomially transmitted infections in the NICU as have CA infections in previously healthy neonates at Texas Children’s Hospital (TCH). Postpartum women in endemic areas are also at risk for S aureus skin infections, including mastitis, and cesarean incision infections after an uncomplicated pregnancy and delivery.

The increasing prevalence of neonatal and maternal CA S aureus infections in Houston and other major metropolitan areas throughout the United States underlines the importance of rapidly identifying potentially infected neonates and instituting effective therapy. Treatment recommendations for CA S aureus disease in previously healthy neonates are limited. For example, there are no clear guidelines regarding the decision to use local versus systemic antibiotics or oral versus parenteral therapy; empiric antibiotic selection; duration of therapy; inpatient versus outpatient treatment; and the indications for an evaluation for a serious bacterial infection (SBI) obtaining blood, urine, and cerebrospinal fluid (CSF) for culture.

Although multicenter, randomized, controlled trials are needed to determine optimally the safety and efficacy of various protocols, accumulated experience in the evaluation and treatment of these infections in centers with a high prevalence such as TCH is currently available source of guidance. In this report, we describe the evaluation and treatment of CA S aureus disease in neonates who presented to TCH.

METHODS

As previously described, we prospectively identified TCH patients with CA S aureus infections and collected their S aureus isolates from the clinical microbiology laboratory. CA organisms were isolated (1) within 48 hours of hospitalization; (2) during an outpatient visit; or (3) after 48 hours of hospitalization but clinical evidence suggested a CA infection, such as symptoms present at hospital admission. Exclusion criteria were (1) underlying illness predisposing to frequent hospitalizations or medical visits, (2) indwelling catheters or percutaneous medical devices, and (3) hospitalization within the past year excluding birth. The Infectious Disease Research Laboratory coded and froze the S aureus specimens in horse blood at −80°C. Antibiotic susceptibilities to oxacillin, clindamycin, erythromycin, gentamicin, trimethoprim-sulfamethoxazole, and vancomycin were determined by disk diffusion and categorized by Clinical and Laboratory Standards Institute guidelines. Inducible resistance to clindamycin was evaluated by D test. A research nurse recorded demographic and clinical information from medical charts into an electronic database using a standardized form. The Baylor College of Medicine Institutional Review Board approved this study.

Patients who were ≤30 days of age at onset of symptoms and were seen at TCH between August 1, 2001, and July 31, 2006, were selected from the database. We identified previously healthy neonates who were born at ≥36 weeks’ gestation, were treated for S aureus infection, and presented after nursery discharge. “Previously healthy” was defined as no hospitalizations since birth and no surgeries except circumcision. A data collection form was developed to include more detailed information regarding the maternal history, clinical presentation and manifestations, treatment, and outcome. Data were entered into a separate database. This population included 89 neonates whose clinical presentation and isolates were previously described and 37 neonates who were seen in the subsequent 18 months.

Neonates were classified by the type of skin and soft tissue manifestation. Infections extending beyond skin and soft tissue structures were classified as invasive. Lesions that contained pus <1 cm in diameter were considered pustules. Larger pus-filled lesions were classified as abscesses. Pustulosis was considered local when only 1 region of the body was involved versus diffuse when >1 region of the body (eg, groin and arm) was affected (Fig 1). Urinary tract infection (UTI) was defined as >1000 colony-forming units (CFU) of bacteria per mL when obtained from a catheterized specimen, unless other evidence indicated an infection. CSF pleocytosis was defined as white blood cell (WBC) count of >22/µL after decreasing the WBC count by 1 for every 1000 red blood cells per µL. Fever was defined as a temperature ≥38°C by any route. In the emergency center, rectal temperatures were obtained.

![FIGURE 1](image-url)  
**FIGURE 1**  
Late-preterm, previously healthy male neonate with CA S aureus localized groin pustulosis.
Pulsed-Field Gel Electrophoresis (PFGE)
PFGE was performed for isolates that infected neonates with CSF pleocytosis, and strain relationships were determined as previously described.12–14

Statistical Analysis
Statistical analysis was performed using Fisher’s exact test, $\chi^2$, or $\chi^2$ for trend for dichotomous variables or Student’s $t$ test for comparison of means by using GraphPad Prism 4.0 (GraphPad, San Diego, CA). A Welch’s $t$ test for comparison of means by using GraphPad Prism 4.0 was applied for populations with unequal variance. Analyses were 2-tailed, and $P<.05$ was considered significant.

RESULTS
Population characteristics are described in Table 1. Symptoms began earlier for patients with pustulosis than for those with cellulitis/abscess or invasive infections ($P < .01$). Fever at presentation was more likely as the severity of disease increased (pustulosis versus cellulitis/abscess: $P < .01$; cellulitis/abscess versus invasive: $P = .02$; pustulosis versus invasive: $P < .01$). The admission status and treatment for different disease manifestations are described in Fig 2.

Pretreatment
Known antibacterial treatments before hospital presentation included topical ($n = 8$), oral ($n = 9$), and intramuscular ceftriaxone ($n = 1$) and oral and topical combinations ($n = 3$). Topical medications included mupirocin, bacitracin, Neosporin, and erythromycin eye ointment. Amoxicillin, amoxicillin/clavulanate, azithromycin, cephalexin, and clindamycin were oral medications used. One patient received mupirocin solely for UTI prophylaxis. The child who was treated the longest before hospital presentation was initially prescribed 7 days of oral clindamycin for a diaper area pustulosis without culture, but treatment was extended to 10 days for a recurrence after the last dose. He was admitted with mastitis that required drainage, and the infecting isolate expressed constitutive clindamycin resistance.

Disposition
After evaluation, 15 (12%) patients were discharged from TCH to continue outpatient treatment, and the remaining patients were admitted (Fig 2). No patient who was discharged from the emergency center was febrile or underwent an SBI evaluation. Two children were treated before presentation with oral antibiotics (1 for UTI prophylaxis), and 1 child was treated with topical antibiotic therapy. Ten children were treated with topical antibiotics alone. Five children were treated with oral antibiotics with or without topical medications for abdominal/groin area infections. One infant with a perirectal abscess underwent incision and drainage in the emergency center before discharge. Documented duration of prescribed outpatient treatment ranged from 7 to 14 days ($n = 4$) for topical and 7 to 12 days ($n = 4$) for oral antibiotic therapy. Two additional patients were discharged from the hospital without antibiotics. One patient was recalled for admission because of a positive skin-culture result. One patient initially presented with decreased movement and pain at the right shoulder and temperature of 100°F. He had worsening symptoms and was subsequently admitted with osteomyelitis and septic arthritis.

Initial Evaluation
All cases with the exception of 3 invasive cases were evaluated by culture and Gram stain of the material from the primary site of infection. The Gram stain revealed Gram-positive cocci for 79 (65%) of 121 patients. Potential pathogens other than $S$ aureus that were isolated from primary infection site cultures included $Klebsiella$ species ($n = 12$), $Escherichia coli$ ($n = 9$), $Enterobacter$ species ($n = 6$), $Clostridium$ species ($n = 2$), $Enterococcus$ species ($n = 2$), $Serratia$ species ($n = 1$), $Proteus$ species ($n = 1$), $Pseudomonas$ species ($n = 1$), $Pantoea$ species ($n = 1$), $Acinetobacter$ species ($n = 1$), $Bacillus$ species ($n = 1$), and $Propionibacterium$ species ($n = 1$). Cultures from patients with drainage procedures contained other organisms less frequently than those from patients without drainage procedures (9 of 49 [18%] vs 29 of 77 [38%]: $P = .03$). Four infants received antibiotic coverage for bacterial co-infection with $Klebsiella$ species or $Enterobacter$ species at the primary $S$ aureus infection site. There was no relationship between the initial WBC count or absolute neutrophil count and the likelihood of a diagnosis of invasive disease. When the percentage of immature neutrophils was compared with total neutrophil count, patients with invasive infections had increased ratios compared with patients with pustulosis and cellulitis/abscess; patients with cellulitis/abscess had increased ratios compared with patients with pustulosis; however, the sensitivity for invasive infection using a threshold of 10% was only 67%.

Many neonates underwent evaluation for SBI (Table 1). One uncircumcised boy with phimosis had 10 000 CFU/mL $Enterobacter$ species isolated from the urine; no initial urinalysis was obtained. His groin pustulosis included penile lesions, and his pustular culture grew both $S$ aureus and $Enterobacter$ species. Another uncircumcised boy had a urine culture with only 1000 CFU/mL $S$ aureus but subsequently had purulent material aspirated from a perinephric abscess, which also grew MRSA. A third circumcision boy had >100 000 CFU/mL $S$ aureus isolated from the urine in addition to $S$ aureus isolated from synovial fluid associated with septic arthritis; the initial urinalysis revealed 11 to 20 WBCs per high-power field. Two additional boys were treated but did not meet the UTI criteria.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Pustulosis</th>
<th>Cellulitis/Abcess</th>
<th>Invasive</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population, n</td>
<td>43</td>
<td>68</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>37 (86)</td>
<td>43 (63)</td>
<td>11 (73)</td>
<td>.01</td>
</tr>
<tr>
<td>Symptom onset, days of age</td>
<td>10.6 ± 5.0</td>
<td>15.1 ± 7.1</td>
<td>16.6 ± 7.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Maternal history of SSTI, n (%)</td>
<td>8 (19)</td>
<td>9 (13)</td>
<td>3 (20)</td>
<td>NS</td>
</tr>
<tr>
<td>Family history of SSTI, n (%)</td>
<td>11 (26)</td>
<td>13 (19)</td>
<td>7 (47)</td>
<td>.04</td>
</tr>
<tr>
<td>Breastfeeding, n (%)</td>
<td>22/39 (56)</td>
<td>32/55 (58)</td>
<td>7/14 (50)</td>
<td>NS</td>
</tr>
<tr>
<td>Circumcision, n (%)</td>
<td>25/33 (76)</td>
<td>22/37 (59)</td>
<td>5/10 (50)</td>
<td>NS</td>
</tr>
<tr>
<td>Vaginal delivery, n (%)</td>
<td>27/39 (69)</td>
<td>36/62 (58)</td>
<td>9/15 (60)</td>
<td>NS</td>
</tr>
<tr>
<td>Fever at or before presentation, n (%)</td>
<td>1 (2)</td>
<td>17 (25)</td>
<td>9 (60)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Susceptibility testing, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRSA</td>
<td>30 (70)</td>
<td>45 (66)</td>
<td>9 (60)</td>
<td>NS</td>
</tr>
<tr>
<td>Clindamycin resistant</td>
<td>4 (9)</td>
<td>8 (12)</td>
<td>1 (7)</td>
<td>NS</td>
</tr>
<tr>
<td>Evaluation, n (%)</td>
<td>44</td>
<td>71</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Gram stain with Gram-positive cocci</td>
<td>24/43 (56)</td>
<td>49/70 (70)</td>
<td>6/8 (75)</td>
<td>NS</td>
</tr>
<tr>
<td>Urine culture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive for S aureus</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (29)</td>
<td>.003</td>
</tr>
<tr>
<td>Obtained</td>
<td>22 (50)</td>
<td>50 (70)</td>
<td>7 (64)</td>
<td>.03</td>
</tr>
<tr>
<td>Blood culture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive for S aureus</td>
<td>0 (0)</td>
<td>2 (3)</td>
<td>4 (36)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Obtained</td>
<td>31 (70)</td>
<td>60 (85)</td>
<td>11 (100)</td>
<td>.01</td>
</tr>
<tr>
<td>CSF culture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive for S aureus</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>NS</td>
</tr>
<tr>
<td>Obtained</td>
<td>24 (55)</td>
<td>50 (70)</td>
<td>10 (91)</td>
<td>.01</td>
</tr>
<tr>
<td>Drainage procedure</td>
<td>0 (0)</td>
<td>41 (58)</td>
<td>8 (73)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Empiric intravenous treatment, n (%)</td>
<td>31</td>
<td>67</td>
<td>13</td>
<td>.04</td>
</tr>
<tr>
<td>No. of antistaphylococcal antibiotics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4 (13)</td>
<td>2 (3)</td>
<td>3 (23)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>24 (77)</td>
<td>53 (79)</td>
<td>7 (54)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3 (10)</td>
<td>11 (16)</td>
<td>3 (23)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0 (0)</td>
<td>1 (1)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Vancomycin</td>
<td>6 (19)</td>
<td>31 (46)</td>
<td>9 (69)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Clindamycin</td>
<td>21 (68)</td>
<td>41 (61)</td>
<td>2 (15)</td>
<td>.007</td>
</tr>
<tr>
<td>Nafcillin</td>
<td>3 (10)</td>
<td>6 (9)</td>
<td>2 (15)</td>
<td>NS</td>
</tr>
<tr>
<td>Gentamicin</td>
<td>22 (71)</td>
<td>49 (73)</td>
<td>10 (77)</td>
<td>NS</td>
</tr>
<tr>
<td>Second- or third-generation Cephalosporin</td>
<td>2 (6)</td>
<td>8 (12)</td>
<td>6 (46)</td>
<td>.003</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>13 (42)</td>
<td>34 (51)</td>
<td>11 (85)</td>
<td>.02</td>
</tr>
<tr>
<td>Acyclovir</td>
<td>6 (19)</td>
<td>2 (3)</td>
<td>0 (0)</td>
<td>.005</td>
</tr>
<tr>
<td>Definitive intravenous treatment, n (%)</td>
<td>17</td>
<td>43</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>MRSA</td>
<td></td>
<td></td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>Clindamycin</td>
<td>11 (92)</td>
<td>26 (84)</td>
<td>6 (75)</td>
<td></td>
</tr>
<tr>
<td>MSSA</td>
<td></td>
<td></td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>Clindamycin</td>
<td>2 (40)</td>
<td>7 (64)</td>
<td>1 (20)</td>
<td></td>
</tr>
<tr>
<td>Nafcillin</td>
<td>3 (60)</td>
<td>4 (36)</td>
<td>4 (80)</td>
<td></td>
</tr>
<tr>
<td>Central venous line placement</td>
<td>1 (6)</td>
<td>9 (21)</td>
<td>5 (36)</td>
<td>.05</td>
</tr>
<tr>
<td>Systemic therapy length, n</td>
<td>30</td>
<td>58</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Oral alone, mean ± SD (n)</td>
<td>11.0 ± 1.4 (2)</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Intravenous and oral, mean ± SD (n)</td>
<td>4.3 ± 1.5 (16)</td>
<td>5.1 ± 2.3 (31)</td>
<td>15 (1)</td>
<td>NS</td>
</tr>
<tr>
<td>Oral</td>
<td>7.0 ± 2.6 (16)</td>
<td>7.0 ± 2.0 (31)</td>
<td>7 (1)</td>
<td>NS</td>
</tr>
<tr>
<td>Total</td>
<td>11.25 ± 2.5 (16)</td>
<td>12.1 ± 2.3 (31)</td>
<td>22 (1)</td>
<td>NS</td>
</tr>
<tr>
<td>Intravenous alone, mean ± SD (n)</td>
<td>7.0 ± 2.6 (12)</td>
<td>9.6 ± 3.7 (27)</td>
<td>20.2 ± 12.2 (13)</td>
<td>.03, .001, &lt;.001</td>
</tr>
</tbody>
</table>

SSTI indicates skin or soft tissue infection; NS, not significant.

* Asterisk indicates skin or soft tissue infection; NS, not significant.

** Pustulosis versus cellulitis/abscess.

*** Pustulosis versus invasive.

**** Cellulitis/abscess versus invasive.

† Trend.

‡ Does not include an umbilical venous catheter in the patient who died.

§ Two re-admitted patients were analyzed according to planned therapy (focal motor seizures and Flavimonas species central line infection).

∥ Oral alone versus intravenous alone for pustulosis.

### Notes:
- Intravenous alone versus intravenous and oral for cellulitis/abscess.
- Intravenous alone versus intravenous and oral for cellulitis/abscess.
CSF Pleocytosis

None of 84 CSF cultures isolated \(S\) aureus. Eleven infants were found to have CSF pleocytosis (Table 2). Disease severity, season, ethnicity, gender, age, and the presence of bacteremia or UTI were not related to the CSF WBC count. Two infants were treated for the CSF pleocytosis despite negative culture results because they received an antibiotic before their lumbar punctures were performed. Only 4 (36%) patients had a history of fever at the time of presentation. Only 1 patient had a CSF viral evaluation; both the viral culture and herpes simplex virus (HSV) polymerase chain reaction were negative.

TABLE 2  CSF Results for Term and Late-Preterm, Previously Healthy Neonates With CA \(S\) aureus Infection and CSF Pleocytosis

<table>
<thead>
<tr>
<th>Age, d</th>
<th>Gender</th>
<th>Infection</th>
<th>Isolate</th>
<th>Pretreat</th>
<th>Fever</th>
<th>Gram Stain</th>
<th>WBCs, (\text{per mm}^3)</th>
<th>PMN, %</th>
<th>Mono, %</th>
<th>RBCs, (\text{per mm}^3)</th>
<th>Protein, mg/dL</th>
<th>Glucose, mg/dL</th>
<th>aWBC, (\text{per mm}^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>M</td>
<td>Mastitis</td>
<td>MRSA, not tested</td>
<td>No</td>
<td>No</td>
<td>Negative</td>
<td>40</td>
<td>62</td>
<td>36</td>
<td>14,000</td>
<td>91</td>
<td>49</td>
<td>26</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>Submental and groin cellulitis/abscess</td>
<td>MRSA USA300</td>
<td>Yes</td>
<td>Yes</td>
<td>Negative</td>
<td>32</td>
<td>8</td>
<td>92</td>
<td>134</td>
<td>80</td>
<td>102</td>
<td>44</td>
</tr>
<tr>
<td>21</td>
<td>M</td>
<td>Groin abscess</td>
<td>MRSA USA300</td>
<td>Yes</td>
<td>Yes</td>
<td>Negative</td>
<td>50</td>
<td>10</td>
<td>90</td>
<td>8450</td>
<td>NA</td>
<td>NA</td>
<td>42</td>
</tr>
<tr>
<td>24</td>
<td>M</td>
<td>UTI, osteomyelitis, and septic arthritis</td>
<td>MSSA unique</td>
<td>No</td>
<td>No</td>
<td>Negative</td>
<td>44</td>
<td>11</td>
<td>89</td>
<td>2400</td>
<td>102</td>
<td>45</td>
<td>42</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>Groin pustulosis</td>
<td>MRSA USA300</td>
<td>No</td>
<td>No</td>
<td>Negative</td>
<td>44</td>
<td>0</td>
<td>100</td>
<td>47</td>
<td>111</td>
<td>45</td>
<td>44</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>Trunk cellulitis/abscess and early septic shock</td>
<td>MRSA USA300</td>
<td>No</td>
<td>Yes</td>
<td>Negative</td>
<td>218</td>
<td>28</td>
<td>71</td>
<td>137,400</td>
<td>93</td>
<td>94</td>
<td>81</td>
</tr>
<tr>
<td>20</td>
<td>F</td>
<td>Omphalitis and thigh pustule and conjunctivitis</td>
<td>MSSA unique</td>
<td>No</td>
<td>No</td>
<td>Negative</td>
<td>192</td>
<td>30</td>
<td>70</td>
<td>102,250</td>
<td>252</td>
<td>48</td>
<td>90</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>Groin and sacral cellulitis/abscess*</td>
<td>MRSA USA300</td>
<td>No</td>
<td>Yes</td>
<td>Negative</td>
<td>190</td>
<td>54</td>
<td>46</td>
<td>79,937</td>
<td>128</td>
<td>42</td>
<td>110</td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>Abdomen, groin, leg pustulosis</td>
<td>MSSA unique</td>
<td>No</td>
<td>Yes</td>
<td>Negative</td>
<td>183</td>
<td>15</td>
<td>85</td>
<td>103</td>
<td>106</td>
<td>50</td>
<td>183</td>
</tr>
<tr>
<td>25</td>
<td>M</td>
<td>Mastitis</td>
<td>MRSA USA300</td>
<td>No</td>
<td>No</td>
<td>Negative</td>
<td>405</td>
<td>30</td>
<td>65</td>
<td>245</td>
<td>72</td>
<td>49</td>
<td>405</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>Abdomen and groin pustulosis</td>
<td>MRSA USA300</td>
<td>No</td>
<td>No</td>
<td>Negative</td>
<td>825</td>
<td>47</td>
<td>47</td>
<td>72,875</td>
<td>222</td>
<td>56</td>
<td>752</td>
</tr>
</tbody>
</table>

Pretreat indicates systemic antibiotics before lumbar puncture; PMN, polymorphonuclear cells; Mono, mononuclear cells; RBCs, red blood cells; aWBC, adjusted WBC count. M, male; F, female; NA, not available.

* Only child with a repeat lumbar puncture (WBC count of 26 per \(\text{mm}^3\); 3\% PMNs; 97\% mononuclear cells, red blood cell count of 530 per \(\text{mm}^3\)).
Seven (70%) of 10 available isolates from the primary infection site had a PFGE pattern consistent with USA300.

Antibiotic Selection
Intravenous antibiotic selection is described in Table 1. Empiric antibiotics were initiated within 2 days of treatment onset. Definitive antibiotics were the final antistaphylococcal coverage chosen in patients who were hospitalized for >4 days at TCH.

The predominant antistaphylococcal monotherapy empiric choice for pustulosis and cellulitis/abscess was clindamycin for 75% (18 of 24) and 58% (31 of 53), respectively. All patients who had invasive manifestations and were empirically treated for *S aureus* with 1 exception received vancomycin; 33% also received nafcillin or clindamycin. Eighty-nine (71%) patients were treated with gentamicin, ampicillin, or cephalosporins for synergy or empiric coverage for other bacteria. The presence of documented fever had no specific relationship to antibiotic choice. Patients with cellulitis/abscess only in the upper body were more likely to be treated with monotherapy than patients with lower body involvement (94% [32 of 34] vs 64% [21 of 33]; *P* = .003).

Early Discharge Versus Full Inpatient Therapy
Sixty-two patients were discharged after inpatient therapy to complete treatment with oral and/or topical antibiotics (ie, early discharge; Fig 2). Patients with pustulosis and cellulitis/abscess were more likely to have early discharge with oral or topical treatment than patients with invasive infections (*P* < .001 and *P* = .002, respectively). Only 1 patient with invasive disease was discharged on oral antibiotic therapy. This patient had septic shock and trunk cellulitis/abscess and completed 15 days of inpatient therapy before discharge with 7 days of clindamycin. Patients with pustulosis were more likely to be treated with topical therapy only after early discharge than were patients with cellulitis/abscess (*P* = .04). Eighty percent (16 of 20) of patients with mastitis alone completed treatment with outpatient oral antibiotics. Four inpatients (3 with pustulosis and 1 with cellulitis/abscess) received topical antibiotics as their sole inpatient antistaphylococcal antibiotic. Some patients had additional diagnoses that affected the location of care and duration of hospitalization.

Oral Antibiotic Choice After Inpatient Therapy
Clindamycin was the most commonly used oral antibiotic for both MRSA (32 of 33 [97%]) and methicillin-susceptible *S aureus* (MSSA; 8 of 16 [50%]) infection. Five patients were treated with cephalaxin, and 2 patients were treated with amoxicillin/clavulanate. Three patients who were older than 21 days at discharge were treated with trimethoprim-sulfamethoxazole.

Treatment Length
Table 1 describes the length of effective treatment for patients who received antistaphylococcal systemic therapy. Pretreatment with oral antibiotic therapy is not included because the therapy was presumed to be ineffective. Patients who were transferred to other institutions and 1 patient who died were not included in Table 1.

Outcome
Only 1 patient is known to have had a treatment failure after evaluation at TCH. He presented with mastitis and chest cellulitis that measured $2.5 \times 2$ cm at 18 days of age to the TCH emergency center. He received cephalexin before his initial presentation to TCH. Blood, urine, and CSF cultures were negative. Ampicillin, gentamicin, and vancomycin were initiated. He was transferred to another facility by family request. After 72 hours of intravenous ampicillin, gentamicin, and clindamycin and an attempted aspiration, the patient was discharged on oral clindamycin. He returned to TCH 4 days after discharge, at 25 days of age, with a $5 \times 4$ cm mastitis and was admitted on intravenous clindamycin. The next day, a mastotomy was performed with removal of 8 mL of purulent material that grew MRSA that was susceptible to clindamycin. He was discharged at 31 days of age after completing his entire intravenous antibiotic course.

Five patients had a second *S aureus* infection after an initial neonatal CA *S aureus* infection. All infections occurred >1 month after the final treatment day and therefore are considered reinfection, not treatment failure. Neonate 1 had an initial MRSA omphalitis and presented at 9 months of age with MRSA osteomyelitis of the femur and distal thumb complicated by UTI, bacteremia, CSF pleocytosis (11 WBCs per mm$^3$), and pericardial effusion. Neonate 2, with MRSA mastitis, developed another cellulitis (no culture obtained) at 5 months of age and an MRSA hand infection at 2 years of age. Neonate 3 initially presented with MSSA groin pustulosis and developed an MSSA otitis externa, otitis media, and early mastoiditis at 7 months of age. Neonate 4 had MRSA mastitis and at 2 months of age developed an MRSA dacryocystitis and preseptal cellulitis. Neonate 5 had an MRSA sacral abscess as a neonate and developed an MRSA scalp abscess at 9 months of age. Other children may have experienced a second infection or treatment failure but were treated elsewhere. For 3 patients, the isolates from the first and second infections had identical susceptibility patterns. Neonate 2 had an initial isolate that was resistant whereas the second isolate was susceptible to erythromycin. Neonate 5 had an initial isolate that was constitutively resistant to clindamycin, whereas the second isolate was susceptible to clindamycin. One additional patient with an extensive family history of skin and soft tissue infections initially had diffuse MRSA pustulosis involving the ears, axilla, and
a tight phimosis. Therefore, term or late-preterm neonates with an unremarkable medical/surgical history, a localized area of pustulosis, no systemic symptoms, and no fever may not require SBI evaluation. A blood culture alone is probably all that is required for neonates who meet these criteria but with diffuse pustulosis involving multiple skin regions. We recommend that all neonates with infections that are more severe than diffuse pustulosis and those with systemic symptoms or fever continue to be evaluated for bacterial infection by blood, urine, and CSF culture until results of prospective studies are available.

In this study, 11 neonates had CSF pleocytosis with negative CSF bacterial culture results. Four infants had CSF pleocytosis with adjusted WBC counts >100/mm³. The cause or pathogenesis of this pleocytosis is unknown. The meningeal inflammation may have a pathogenesis similar to the mononuclear aseptic meningitis described concurrent with UTI in infants. Potential mechanisms include known or as yet unidentified proinflammatory molecules that are produced by S aureus and can cross the intact blood-brain barrier or a concurrent viral meningitis. Because only 1 patient had a CSF evaluation for viral infection that was negative, the possibility of viral meningitis cannot be excluded.

As our experience with CA S aureus infections has increased, treating patients with local pustulosis in the outpatient setting with topical antibiotic therapy alone seems to be reasonable as was preliminarily recommended by some experts. We recommend admitting all neonates with infections that are more severe than pustulosis and those who receive any evaluation for bacterial infection of the blood, urine, or CSF. Empiric antibiotics should be based on the local susceptibility pattern of CA S aureus isolates; clindamycin with or without gentamicin for patients with noninvasive disease and vancomycin, nafcillin, and gentamicin for patients with invasive infections are commonly used in our setting, where <10% of CA S aureus isolates are resistant to clindamycin.

It is essential that criteria for the length of inpatient therapy be established for patients with noninvasive infections who are discharged to continue treatment at home. These criteria should include resolution of systemic symptoms and fever, improvement in the disease, and that any systemic bacterial cultures obtained not isolate pathogens after 48 hours of incubation. Other factors to consider include the availability of antibacterial susceptibility results and parental cooperation to adjust therapy to the most appropriate antibiotic choice after discharge, if necessary. Patients with invasive infections should complete antibiotic treatment intravenously. Some patients who require a prolonged treatment course (eg, osteomyelitis) may be candidates for home intravenous antibiotics.

Forty-nine neonates were treated with oral antibiotics
after initial intravenous therapy. Amoxicillin-clavulanate and cephalixin have adequate oral bioavailability in the neonate.\textsuperscript{19,20} Clindamycin has adequate oral bioavailability when administered to infants 1 to 6 months of age.\textsuperscript{21} Although similar information is not available for neonates, clindamycin is approved for oral administration in neonatal infection.\textsuperscript{22} Because trimethoprim-sulfamethoxazole may increase the risk for kernicterus by displacing bilirubin in the first few weeks of age, its use should be delayed beyond the immediate neonatal period.\textsuperscript{23–25} This issue may be important in areas of increasing clindamycin resistance. Linezolid is another oral antibiotic alternative that is approved for neonatal use\textsuperscript{26} and may be considered when the \textit{S. aureus} isolate is resistant to other oral antibiotic choices. This medication should be given under close supervision by an infectious disease expert.

We found only 1 patient with an outpatient treatment failure on oral antibiotics. Possible explanations for his treatment failure include poor antibiotic compliance, poor oral antibiotic absorption with inadequate levels to penetrate the abscess site, and the need for adequate drainage. Other children may have experienced outpatient treatment failure but were treated by a non-TCH health care provider and the chart was not available to us. A comprehensive follow-up program for neonates is imperative when treatment is initiated or completed in the outpatient setting to monitor carefully the response to therapy.

Clinical studies are needed to determine duration of therapy for noninvasive infections, because subjective durations of 5, 7, or 10 days are often prescribed. As highlighted in Table 1, the average oral prescription length after intravenous antibiotics was 7 days for both patients with pustulosis and patients with cellulitis/abscess despite clearly different disease severity.

**CONCLUSIONS**

Practitioners at our institution are using varied evaluation and treatment strategies for neonatal CA \textit{S. aureus} disease, and prospective trials with close follow-up are needed to determine the optimal management strategy, including duration of treatment based on response to therapy in this vulnerable population.

**ACKNOWLEDGMENTS**

This study was supported in part by a grant from Pfizer, Inc.

We thank Linda B. Lamberth for technical assistance.

**REFERENCES**

VALUE-BASED PURCHASING: PUTTING PAY ON THE LINE TO IMPROVE HEALTH CARE

“For many doctors there are no consequences for poor performance—no financial penalties from the insurers who pay most of the bills. Unbeknownst to many busy physicians, this is all changing. Under a federal law that took effect this summer, Medicare will pay doctors a 1.5 percent bonus incentive for reporting quality measures set out by a program called P.Q.R.I., for Physician Quality Reporting Initiative. This represents a significant step toward public reporting and pay for performance. Hospital quality data is already being posted on Medicare’s Web site, www.hospitalcompare.hhs.gov. It is now easy to find out how a given hospital compares with the national average for death rates from common illnesses like heart attack, as well as the hospital’s performance on 21 quality measures (like giving aspirin and smoking-cessation advice to heart attack patients). Hospitals now receive incentives for reporting such data; soon their actual payments will be based on their performance on these measures. Just as managed care swept the health system a decade ago and left its mark, so should this new policy, which Medicare calls value-based purchasing. Soon, doctors will be paid according to quality and value. And our performance will be publicly reported on the Internet so our patients can choose the best doctors. With doctors’ reputations at stake, they will have an incentive to improve. Though value-based purchasing sounds promising, such an approach has several potential land mines. Doctors will have an incentive to treat the less ill patients, because the outcome is bound to be better. Imagine the consequences if doctors cherry-pick patients on that basis. (Medicare does adjust its report cards for severity of illness, yet doctors are skeptical.) Also, doctors already fed up with the bureaucracy and paperwork of Medicare may wash their hands of the program and refuse to see any Medicare patients, just as vast numbers of baby boomers begin turning 65 and entering the program. How value-based purchasing will affect health care only time and human behavior will tell. But change is urgent for our misaligned health care system.”


Noted by JFL
Are We Overprescribing Antireflux Medications for Infants With Regurgitation?

Vikram Khoshoo, MD, PhD, Dean Edell, MD, MPH, Aaron Thompson, MD, Mitchell Rubin, MD

Pediatric Specialty Center, West Jefferson Medical Center, New Orleans, Louisiana

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Our goal was to evaluate the diagnosis and treatment of infants with persistent regurgitation who were referred to a pediatric gastroenterology service.

METHODS. The records of 64 infants with persistent regurgitation and without any neurodevelopmental abnormalities, underlying illness, or cigarette smoke exposure were evaluated for diagnostic workup and treatment. Forty-four infants underwent extended esophageal pH monitoring.

RESULTS. Only 8 of 44 pH studies showed abnormal acid reflux. Forty-two of these 44 infants were already on antireflux medications. Other etiologies included hypertrophic pyloric stenosis (4) and renal tubular acidosis (1). Discontinuation of medication did not result in worsening of symptoms in most infants with normal pH studies.

CONCLUSIONS. The majority of infants who were prescribed antireflux drugs did not meet diagnostic criteria for gastroesophageal reflux disease.
DURING 1998–1999, INFANTS with regurgitation ac-
counted for 14% of our total patient referrals, and
40% of these infants were already on anti-gastroesop-
phageal reflux disease (GERD) medications and/or an ex-
tensively hydrolyzed formula. During 2006–2007, such
infants accounted for 23% of all referrals, and 90% of 
them were already receiving anti-GERD medications 
and/or extensively hydrolyzed formulas. This shift 
prompted us to evaluate the diagnosis and treatment of 
infants with regurgitation.

METHODS
We evaluated the medical charts of infants who were 
referred to the Pediatric Specialty Center at West Jeffer-
son Medical Center during a 3-year period for symptoms 
of regurgitation lasting for >2 weeks. The following 
groups of infants were excluded: (1) those who were 
born preterm or small for gestational age; (2) those with 
any underlying illness, especially diarrhea; (3) those 
with neurodevelopmental abnormalities or dysmorphic 
features; (4) those with a lower respiratory tract illness;
(5) those who were exposed directly or indirectly to 
cigarette smoke; and (6) those whose care was provided 
by anyone other than the mother.

For the sake of studying a homogeneous population 
with comparable growth patterns, energy intake and 
maturity, we included only those infants who were born 
at term, appropriate for gestational age, and of normal 
neurodevelopmental status. Cigarette smoke exposure 
and lower respiratory tract illnesses, especially asthma, 
can exacerbate reflux, so such children were excluded to 
maintain consistency in clinical status. For achieving 
reliability in assessing oral intake and formula prepara-
tion, infants who were provided care by caregivers other 
than their mother were excluded. Chart reviews and 
data collection were performed after approval was ob-
tained from the institutional review board. The data 
were gathered on a standardized form designed in our 
clinic and used to collect information on all patients as 
part of our ongoing research on GERD.

The diagnostic tests that were performed either before 
or after referral were radiology, ultrasonography, ex-
tended esophageal pH monitoring, blood chemistry, 
and blood gas analysis. The pH study was performed after 
withdrawal of the prokinetics or H2-receptor antagonists 
for at least 5 days and proton-pump inhibitors for at least 
10 days. A pH of <4 in the distal esophagus for >5% of 
the time was regarded as an abnormal study and a 
marker of GERD. The upper gastrointestinal barium 
studies were used only for diagnosing structural abnor-
malities, and any mention of reflux by the radiologist 
was not considered diagnostic of GERD.

RESULTS
Ninety-two infants were identified, and 64 fulfilled all 
entry criteria. The main reasons for exclusion were care-
givers other than the mother (n = 8), shorter duration of 
symptoms (n = 3), exposure to cigarette smoke (n = 3), 
underlying illness (n = 3), prematurity (n = 2), and 
multiple factors (n = 9). We received referrals from 132 
physicians, and 42 of them accounted for referring 84% 
of the infants included in this study. The clinical presenta-
tion, patient details, and results of diagnostic tests are 
given in Table 1.

Most (89%) infants were gaining weight at a rate of 
>15 g/day. Feeding >504 kJ/kg per day (>120 kcal/kg per 
day) was noted in 40.6% infants. Thickening of feeds 
(81.3%) and antireflux medication (90.6%) were com-
mon interventions used for control of symptoms before 
referral. Dry rice cereal (84.6%) and dry oatmeal cereal 
(15.4%) were the 2 thickening agents used. Inappropri-
ate thickening with <1 tablespoon of cereal per ounce of 
formula was very common (70.3%). In 85.9% of the 
infants, the formula had been changed in the 2 weeks 
before the referral. A completely or extensively hydro-
lyzed formula was commonly used (60.9%).

Diagnostic tests yielded abnormal results in 13 
(20.3%) infants and were as follows: (1) GERD (n = 8)

<table>
<thead>
<tr>
<th>TABLE 1 Patient Details and Results of Diagnostic Workup for Infants Who Were Referred for Suspected GERD (N = 64)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Age, mean ± SD, wk</td>
</tr>
<tr>
<td>Gender, male/female</td>
</tr>
<tr>
<td>Duration of symptoms, mean ± SD, wk</td>
</tr>
<tr>
<td>Symptoms, No. (%) of patients</td>
</tr>
<tr>
<td>Regurgitation alone</td>
</tr>
<tr>
<td>Regurgitation with irritability</td>
</tr>
<tr>
<td>Regurgitation with feeding problems</td>
</tr>
<tr>
<td>Regurgitation with respiratory symptoms</td>
</tr>
<tr>
<td>Mean weight gain for past 4 wk, No. (%) of patients</td>
</tr>
<tr>
<td>&gt;15 g/d</td>
</tr>
<tr>
<td>&lt;15 g/d</td>
</tr>
<tr>
<td>Formula change in the past 2 wk, No. (%) of patients</td>
</tr>
<tr>
<td>Soy based</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Thickenfed feeds, No. (%) of patients</td>
</tr>
<tr>
<td>Appropriate concentration</td>
</tr>
<tr>
<td>Inappropriate concentration</td>
</tr>
<tr>
<td>Energy intake, No. (%) of patients</td>
</tr>
<tr>
<td>&gt;504 kJ/kg per d (&gt;120 kcal/kg per d)</td>
</tr>
<tr>
<td>378–504 kJ/kg per d (90–120 kcal/kg per d)</td>
</tr>
<tr>
<td>&lt;378 kJ/kg per d (&lt;90 kcal/kg per d)</td>
</tr>
<tr>
<td>Anti-GERD medications, No. (%) of patients</td>
</tr>
<tr>
<td>H2-receptor antagonists</td>
</tr>
<tr>
<td>H2-receptor antagonist plus metoclopramide</td>
</tr>
<tr>
<td>Metoclopramide</td>
</tr>
<tr>
<td>Proton-pump inhibitors</td>
</tr>
<tr>
<td>Proton-pump inhibitor plus metoclopramide</td>
</tr>
<tr>
<td>Diagnostic tests, No. of patients (abnormal)</td>
</tr>
<tr>
<td>Upper gastrointestinal imaging</td>
</tr>
<tr>
<td>Ultrasonography</td>
</tr>
<tr>
<td>Blood gas analysis</td>
</tr>
<tr>
<td>Extended esophageal pH monitoring</td>
</tr>
</tbody>
</table>

PEDIATRICS Volume 120, Number 5, November 2007 947
based on abnormal esophageal pH with or without impedance studies; (2) hypertrophic pyloric stenosis \((n = 4)\) based on ultrasonography; (3) metabolic alkalosis \((n = 4)\) and metabolic acidosis \((n = 1)\) based on blood gas analysis; the 4 infants with metabolic alkalosis had hypertrophic pyloric stenosis, and the 1 with metabolic acidosis had renal tubular acidosis; and (4) gastric outlet obstruction \((n = 2)\) based on radiology; both had hypertrophic pyloric stenosis.

A total of 44 extended esophageal pH/impedance monitoring studies were performed on 44 infants. Forty-two of these 44 infants were already on anti-GERD medications. The results of only 8 \((18.2\%)\) of 44 of these studies were abnormal. Of the remaining 20 infants who did not get a pH-probe study, 4 had pyloric stenosis and 1 had renal tubular acidosis. The 15 remaining infants were thriving and had only regurgitation without any accompanying symptoms and were expected to have a low probability for an abnormal pH-probe study and so did not get one.

Regurgitation alone as the presenting symptom was more common in infants with a normal pH-probe study \((69\%)\) as compared with those with an abnormal pH-probe study \((25\%)\). Seven infants had weight gains of \(<15\) g/day, and these included the 4 infants with pyloric stenosis, 1 with renal tubular acidosis, and 2 others with abnormal pH-probe study results.

As part of our standard of care, follow-up of all patients is through clinic visits, e-mails, and telephone calls. All infants included in this study had a follow-up scheduled for 1 to 4 weeks after their evaluation. The 8 infants with abnormal pH-probe–study results and 1 infant with renal tubular acidosis stayed on treatment with anti-GERD medications. Four infants underwent pyloromyotomy for hypertrophic pyloric stenosis. After surgery, 1 infant became asymptomatic and did not require any medication. Of the remaining 3 infants, 2 had persistent regurgitation postoperatively and were treated with metoclopramide and unthickened feeds for presumed delayed gastric emptying, and 1 infant with regurgitation and irritability was treated with unthickened feeds along with metoclopramide and nizatidine for a clinical diagnosis of GERD and delayed gastric emptying.

All anti-GERD medications were withdrawn in the remaining 51 infants, but they continued on their existing formula. The parents of all of these infants were counseled on appropriate thickening and intake of formula as well as positioning. There was no follow-up for 6 of 51 infants, 6 infants showed worsening of symptoms, and 39 had improvement/no worsening of symptoms.

**DISCUSSION**

On the basis of the data of this study, we believe that most infants who had reflux symptoms and were referred to our pediatric specialty service did not meet the strict diagnostic criteria for GERD yet had received some anti-GERD medications. Withdrawal of the medications did not result in any worsening in the majority of these infants. Only \(~20\%)\ of the infants in the study had evidence of an underlying pathology to explain their symptoms, such as GERD, pyloric stenosis, or renal tubular acidosis. This is similar to the findings of Condino et al,\(^1\) who reported the occurrence of normal pH-probe and impedance studies in most infants who were referred for reflux symptoms. Overfeeding and under-thickening of the formula were noted frequently. Feeding intakes of \(>504\) kJ/kg per day \((>120\) kcal/kg per day) were common and could have contributed to the symptoms. Thickening of feeds was commonly advised, but concomitant reduction in the volume was not made, resulting in excessive volume/energy intake. Overfeeding results in gastric distension that is a precipitating factor for inappropriate relaxation of lower esophageal sphincter, a hallmark of reflux. We previously described the efficacy and composition of smaller volume thickened feeds that will reduce the frequency of reflux by \(~50\%).\(^2\) Such feeds need to be small volume and thickened with 1 tablespoon of dry rice cereal per ounce of formula, keeping in mind not only the volume but also the total energy intake for weight and age. Similar findings also were reported by Orenstein et al,\(^3\) who used 1 tablespoon of rice cereal per ounce of formula. In our patients, acid-suppressing drugs were used frequently for reflux symptoms, although no objective diagnosis of GERD was established. The persistence of regurgitation despite anti-GERD medications was no surprise because these drugs are unlikely to reduce the frequency of reflux.\(^4\) Moreover, the simple use of appropriate conservative therapy alone instituted via a telephone triage system can lead to improvement of symptoms in \(24\%)\ of such infants.\(^5\)

It was interesting to note that for infants with abnormal pH-probe results, the presenting complaints often included other symptoms in addition to regurgitation, whereas for infants with normal pH-probe results, regurgitation alone was usually the only presenting symptom. Also, a weight gain of \(<15\) g/day was always associated with an underlying pathology such as pyloric stenosis, renal tubular acidosis, or GERD.

In evaluating our results, 2 concerns should be kept in mind. First, the patients in this study were a select population of infants who had persistence of symptoms and were referred for subspecialist care. Second, there may be regional differences in the way certain conditions are managed by primary care physicians.

This early use of anti-GERD medications in infants with regurgitation could be for several reasons: (1) lack of a simple diagnostic tool for GERD in infants for the primary care physician may prompt the use of a therapeutic trial, (2) parental anxiety at persistence of symptoms may prompt the use of a therapeutic trial, and (3)
aggressive marketing by the manufacturers of acid-suppressing drugs seems to have blurred the dividing line between gastroesophageal reflux and GERD.

CONCLUSIONS
We have shown that the parental perception of the volume of their child’s emesis is greatly exaggerated, usually by fivefold to sixfold (eg, 5 mL of emesis is perceived as 30 mL). Parent–doctor communication, parent education, and reassurance are very important tools in treating such infants. For an infant with persistent regurgitation and no red flags (poor weight gain, excessive crying, irritability, feeding problems, recurrent respiratory symptoms, hoarseness, chronic cough, disturbed sleep, or hematemesis), the physician should be able to persist with conservative treatment (eg, appropriate small-volume thickened feeds, correct positioning, avoidance of cigarette smoke, trial of a hypoallergenic formula) without parental anxiety. The recently simplified and validated Infant Gastroesophageal Reflux Questionnaire–Revised may be a consideration for primary care physicians to use as a fairly accurate and a noninvasive tool for not only diagnosing GERD in infants but also following symptom progression. The North American Society for Pediatric Gastroenterology, Hepatology and Nutrition (www.naspghan.org) has provided a very helpful and comprehensive statement that can assist primary care physicians in treating infants and children with gastroesophageal reflux.

REFERENCES
Burden of Acute Sore Throat and Group A Streptococcal Pharyngitis in School-aged Children and Their Families in Australia

Margaret H. Danchin, PhD\textsuperscript{a,b,c}, Susan Rogers, MSc\textsuperscript{b}, Loraine Kelpie, BHlthSc(Nurs)\textsuperscript{b}, Gowri Selvaraj, BSc\textsuperscript{b}, Nigel Curtis, PhD\textsuperscript{a,b,c},

John B. Carlin, PhD\textsuperscript{a,b}, Terence M. Nolan, PhD\textsuperscript{b,d}, Jonathan R. Carapetis, PhD\textsuperscript{a,b,c,e}

\textsuperscript{a}Department of Paediatrics and \textsuperscript{b}School of Population Health, University of Melbourne, Melbourne, Australia; \textsuperscript{c}Department of General Medicine, Royal Children's Hospital, Melbourne, Australia; \textsuperscript{d}Menzies School of Health Research, Charles Darwin University, Darwin, Australia; \textsuperscript{e}Murdoch Children's Research Institute, Parkville, Australia

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The objective of this study was to determine the incidence, transmission, carriage, and risk factors for group A streptococcal pharyngitis in school-aged children and their families.

METHODS. A 16-month, prospective, family-based cohort study was undertaken from August 2001 through December 2002 in Melbourne, Australia. A total of 202 families (853 people) with at least 1 child aged 3 to 12 years were randomly selected from 3 primary care practices across suburban Melbourne to collect surveillance data for acute group A streptococcal pharyngitis, including serology for index and secondary cases and intermittent carriage data. Cohort retention was 97\% for 16 months.

RESULTS. The incidence of acute sore throat, group A streptococcal swab–positive pharyngitis, and serologically confirmed group A streptococcal pharyngitis was 33, 13, and 8 per 100 child-years, respectively, for school-aged children (5–12 years) and 60, 20, and 15 per 100 family-years, respectively. Sore throat was less common in adults than children, but adults with sore throat were as likely as children to have group A streptococcal culture–positive or serologically proven pharyngitis. In families who had a primary case, 43\% had at least 1 secondary case, and in family members who were at risk, 13\% contracted a secondary case. The spring, summer, and winter carriage rates for children were 13\%, 8\%, and 16\%, respectively, and for adults the rate was 2\% across all seasons.

CONCLUSIONS. Group A streptococcal pharyngitis is still common, and the peak incidence occurs in school-aged children. However, the incidence in adults is higher than expected, and the number of secondary cases in families may be an important factor when considering the potential benefits of treatment.
SORE THROAT REMAINS one of the most common presentations to primary care providers, and group A Streptococcus (GAS) is arguably the only cause of sore throat that needs treatment. The potential aims of treatment are symptom reduction, prevention of transmission, and prevention of supplicative and nonsupplicative streptococcal sequelae, particularly acute rheumatic fever. In developed countries, where the incidence of acute rheumatic fever is low, the role of antibiotics in the treatment of acute pharyngitis is a continuing source of controversy.

Most recent studies of sore throat reported numbers of presentations to primary care, the proportion of sore throats from which GAS can be recovered on throat swabs, or response to antibiotic treatment. Since the 1960s, very few population-based studies have reported the incidence of sore throat or GAS culture–positive sore throat, and no studies have reported the incidence of true GAS pharyngitis (ie, GAS culture–positive sore throat with serologic evidence of recent streptococcal infection). Such data are critical in determining the most appropriate approach to management and in preparing for the arrival of GAS vaccines, which are in clinical trials.

We conducted a prospective study to determine the burden of sore throat, GAS culture–positive sore throat, and serologically proven GAS pharyngitis in suburban Melbourne, Australia. We included all family members in the cohort to assess secondary transmission of GAS within the household.

METHODS

Design
We undertook a 16-month, prospective, family-based cohort study, which was community based, using family medicine practitioners to collect the surveillance data. The Royal Children’s Hospital Ethics in Human Research Committee approved the study.

Setting
Families were resident in 3 diverse geographic and socioeconomic regions of metropolitan Melbourne. Deer Park, Essendon North, and Nepean are in the western, northern, and southern areas of Melbourne, respectively. Deer Park has a diverse mix of ethnicities and a lower socioeconomic status (SES) than the other 2 areas. The median weekly individual income for men in Deer Park was $500 to $599 (Australian dollars) ($255–$305 US dollars) in contrast to Essendon North and Nepean, where the most common weekly individual income for men was $1000 to $1499 (Australian dollars) ($510–$760 US dollars).

Data Collection
Families with at least 1 child aged 3 to 12 years, the peak age of incidence of GAS pharyngitis, were randomly selected and recruited through 3 practices, 1 in each region. Deidentified lists of all families seen within the past 5 to 10 years by the practice were obtained. Initially, 250 families from each practice were selected using random number allocation by computer. Parents were sent a letter signed by the general practitioner, requesting their participation in the study. The remaining families in the practice database were contacted in subsequent mailings, and nonresponders received a second letter until a sufficient number families were recruited. Interested families were contacted, and enrollment visits were arranged. Baseline demographic data were collected, and each family member had a throat swab to determine baseline carriage of GAS.

Families were asked to attend the practice when any member developed a sore throat plus ≥1 of the Centor criteria: a history of fever, tender anterior cervical lymph nodes, pharyngeal exudate, or an absence of cough. Because these criteria were not developed for children, parents were encouraged to bring in their children with a broader set of symptoms, including headache, abdominal pain, vomiting, cough, coryza, and hoarseness. The general practitioner performed a throat swab, recorded the clinical details, and treated the patient as dictated by their usual practice. When the culture isolated GAS, the family was visited at home 1 to 2 weeks later to obtain blood and additional data from the index patient and swabs from family members. The family was visited 2 weeks after the initial blood draw (3–4 weeks after the positive culture) for repeat serology on the index patient, serology on any culture-positive family members, and repeat throat swabs from family members. All participants also had throat swabs taken every 3 or 4 months to assess upper respiratory tract carriage of GAS. To receive a swab at one of these carriage clinics, the participant had to be asymptomatic. Serology was not performed at this time because of logistic (need for paired titers on well individuals) and cost constraints.

Definitions
Children were aged 1 to 18 years, and adults were aged >18 years. The family included all members who were living in the household for the duration of the study period and included all household contacts. A primary case was a symptomatic episode of GAS culture–positive sore throat or serologically confirmed GAS pharyngitis in any family member when all other family members had been asymptomatic for the preceding 2 weeks. A secondary case was an episode of culture-positive or serologically confirmed GAS pharyngitis in a family member that was detected within the 2 weeks of onset of an index patient in the family and was classified as symptomatic or asymptomatic (because all family members had swabs taken once a primary patient was identified in the family). A new episode of pharyngitis was differentiated from a previous episode by a 5-day interval.
asymptomatic period. A relapse of GAS pharyngitis was defined as a recurrent GAS culture–positive episode of sore throat within 30 days of a previous episode with isolation of the same *emm* type at both episodes (discussed later).

Pharyngitis was classified using 3 categories of increasing specificity: all sore throat, culture-positive GAS pharyngitis, and serologically confirmed GAS pharyngitis. Serologically confirmed cases included unequivocal and likely serologically confirmed GAS pharyngitis. Unequivocal GAS pharyngitis was any patient with a positive GAS culture and $>0.2 \log_{10}$ rise in antistreptolysin O (ASO) or anti-DNase B (ADB) titer between acute and convalescent sera. Both titers were performed to increase the chance of detecting evidence of GAS infection, because there is often no clear concordance between the ASO and ADB titers. Likely serologically confirmed GAS pharyngitis included any other case in which the acute serum was taken $>7$ days after the onset of symptoms (by which time a serologic response was already likely to have been mounted$^{12,13}$) and in which the ASO titer and/or ADB titer was greater than the upper limit of normal (ULN) for age. We used ULN values determined recently in Melbourne children$^{14}$ because the ULN values for ASO and ADB titers can vary between populations and within different age groups. When unqualified, the term “serologically confirmed GAS pharyngitis” refers to unequivocal and likely cases together. GAS carriage was defined as the presence of GAS in the upper respiratory tract without an antibody response.$^{15}$ GAS typing, which is mainly based on the M protein, includes serotyping (Lancefield serologic typing) and *emm* sequence typing (genotyping of the amino-terminal portion of the M protein gene).

### Microbiology

All swabs were placed in Amies transport medium, plated within 24 hours on horse blood agar, and incubated at 37°C in 5% CO$_2$ aerobically for up to 24 hours.$^{16,17}$ Beta-hemolytic streptococcal colonies were grouped (A, C, or G) by using latex agglutination (Streptex Kit; Oxoid, Basingstoke, United Kingdom). Senior scientist Wayne Devenish performed the ASO and ADB titers using the routine method of hemolysin neutralization at the Royal Children’s Hospital.$^{18}$

### Sample Size

On the basis of previous estimates, we expected a 10% incidence of GAS culture–positive sore throat per year in our population. A sample size of 160 families ($\sim$240 children), with an intracluster correlation of 0.2, provided an expected width of a standard 95% confidence interval (CI) of $\pm 4\%$ around a 10% point estimate of cumulative incidence and $\pm 5.3\%$ interval around a 20% point estimate. To accommodate up to a 20% dropout rate during the year, we aimed to enroll 200 families, divided evenly among the 3 regions.

### Statistical Analysis

Categorical variables were compared between subgroups by using $\chi^2$ or Fisher’s exact test as appropriate. Univariate Poisson and logistic regression were used to assess potential risk factors for GAS pharyngitis. Data analysis was performed using Stata 7 (Stata Corp, College Station, TX).

### RESULTS

#### Demographics

A total of 202 families were recruited (853 people), 196 families (828 people) of which completed the study, resulting in a retention rate of 97% for both individuals and families (Table 1). Of the 3 regions, Deer Park had the families with the lowest and Nepean had the families with the highest SES and educational status.

#### Sore Throat Episodes

One third of all participants, 16% of adults and 41% of children, reported experiencing a sore throat during the study period (Table 2).

#### Primary and Secondary GAS Culture–Positive Sore Throat

There were 54 primary and 32 secondary cases of GAS culture–positive sore throat (Table 3). Of the sore throat episodes in children, 20% were GAS culture positive compared with 26% in adults, but the proportion of GAS culture–positive sore throat episodes was slightly higher in the 5- to 12-year age group (25%); therefore, although adults were 48% less likely to experience a sore throat compared with children (relative risk [RR]: 0.52 [95% CI: 0.4–0.7]; $P < .0001$), they were just as likely to have a GAS culture–positive throat swab once the sore throat had occurred (RR: 1.2 [95% CI: 0.7–2.0]; $P = .44$).

More than half of the secondary cases were in 5- to 12-year-old children (Table 3). Of the secondary cases, 41% were symptomatic and two thirds occurred in children who were aged 5 to 12 years. Within families who experienced a primary case of GAS pharyngitis, the risk for secondary infection was 1.8 times greater than that of primary infection in the community (RR: 1.8 [95% CI: 1.3–2.4]; $P = .003$).

#### Serologically Confirmed GAS Culture–Positive Pharyngitis

Nineteen primary episodes were unequivocal GAS pharyngitis, and 16 episodes were likely serologically confirmed GAS pharyngitis (total 35 serologically confirmed cases; Table 3). Of the 19 unequivocal primary cases, 15 (32%) had a rise in ASO titer and 10 (21%) had a rise in ADB titer. The geometric mean initial ASO for the unequivocal group (those with a rise in titer) was 350 U.
whereas the mean initial ASO for the likely group (those without a rise in titer) was 733 U. Of the 32 secondary cases, there were 9 unequivocal episodes and 12 likely episodes (total 21 serologically confirmed episodes). There were 3 negative and 8 missing serology episodes. The secondary cases were detected a mean of 8 days after the primary case.

Incidence of Sore Throat, GAS Culture–Positive Pharyngitis, and Serologically Confirmed GAS Pharyngitis

The incidence of acute sore throat, GAS culture–positive pharyngitis, and serologically confirmed GAS pharyngitis were 33, 13, and 8 per 100 person-years, respectively, in children aged 5 to 12 years and 14, 5, and 3 per 100 person-years for adults. The incidence of at least 1 episode of acute sore throat, GAS culture–positive pharyngitis, and serologically confirmed GAS pharyngitis in each family was 60, 20, and 15, respectively, per 100 family-years (Table 2). The estimated incidence of symptomatic GAS culture–positive pharyngitis, including primary and secondary cases, was 10 and 4 per 100 person-years for children aged 5 to 12 years and for adults, respectively (Table 4).

Risk Factors for GAS Culture–Positive Sore Throat

When examined by Poisson regression, both age and season were independent risk factors for GAS pharyngitis. Children experienced more episodes of sore throat and pharyngitis, both culture and serologically confirmed, compared with adults, and there was a bimodal seasonal pattern peaking in winter/spring and autumn. The SES markers of income, education, and location showed no evidence of association with risk for GAS pharyngitis, which was also not associated with gender. There was no clear association with household crowding, the type of facility attended (child care, preschool, or

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Demographic Data for Enrolled Families According to Practice Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td>Participants, n (%)</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>No. of families*</td>
<td>202 (6%)</td>
</tr>
<tr>
<td>No. of people*</td>
<td>877 (49%)</td>
</tr>
<tr>
<td>Adults</td>
<td>393 (35%)</td>
</tr>
<tr>
<td>Male</td>
<td>188 (29%)</td>
</tr>
<tr>
<td>Female</td>
<td>205 (6%)</td>
</tr>
<tr>
<td>Children</td>
<td>484 (14%)</td>
</tr>
<tr>
<td>Age, y</td>
<td>---</td>
</tr>
<tr>
<td>&lt;5</td>
<td>88 (29%)</td>
</tr>
<tr>
<td>5–12</td>
<td>339 (57%)</td>
</tr>
<tr>
<td>13–18</td>
<td>57 (2%)</td>
</tr>
<tr>
<td>Gender</td>
<td>---</td>
</tr>
<tr>
<td>Male</td>
<td>237 (7%)</td>
</tr>
<tr>
<td>Female</td>
<td>247 (7%)</td>
</tr>
<tr>
<td>Facility attended</td>
<td>---</td>
</tr>
<tr>
<td>Day care</td>
<td>37 (2%)</td>
</tr>
<tr>
<td>Preschool/Kindergarten</td>
<td>80 (2%)</td>
</tr>
<tr>
<td>School</td>
<td>327 (10%)</td>
</tr>
<tr>
<td>Total</td>
<td>444 (100%)</td>
</tr>
</tbody>
</table>

* Row percentage is given for totals (families, people); otherwise, percentages provide column-wise breakdown of subgroup categories.

<p>|TABLE 2| Incidence of Primary and Secondary Cases of GAS Pharyngitis According to Age and Family|
|---|---|---|</p>
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Sore Throat</th>
<th>GAS Pharyngitis</th>
<th>Serology-Confirmed Pharyngitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases, n</td>
<td>Time at Risk, Patient-Years</td>
<td>Incidence per 100 Patient-Years</td>
<td>95% CI</td>
</tr>
<tr>
<td>Age, y</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>&lt;5</td>
<td>37 (97)</td>
<td>38.1</td>
<td>26.9–52.6</td>
</tr>
<tr>
<td>5–12</td>
<td>128 (392)</td>
<td>32.7</td>
<td>27.2–38.8</td>
</tr>
<tr>
<td>13–18</td>
<td>26 (65)</td>
<td>40.0</td>
<td>26.1–58.6</td>
</tr>
<tr>
<td>&gt;18</td>
<td>57 (422)</td>
<td>13.5</td>
<td>10.2–17.5</td>
</tr>
<tr>
<td>Total</td>
<td>248 (977)</td>
<td>25.4</td>
<td>22.3–28.7</td>
</tr>
</tbody>
</table>

* Family-years at risk and incidence of at least 1 episode per 100 family-years.
Transmission of GAS Culture–Positive Pharyngitis

Of the families who had a primary case of GAS pharyngitis, 43% had at least 1 secondary case (18 of 42). Of the families who had a secondary case, more than half (11 of 18) had at least 2 secondary cases. Of all people who were at risk (all family members except the index patient), 13% contracted a secondary case (95% CI: 9–18). The emm type of 81% (26 of 32) of the secondary GAS culture–positive isolates was determined, and 96% (25 of 26) of secondary cases in the family had the same emm type as the primary case.

Seasonal Carriage Rates

The GAS carriage rates for children over the 3 seasons were 13%, 8%, and 16% for spring, summer, and winter, respectively. The rate for adults was 2% for each season.

emm Types of Pharyngeal and Carriage Isolates

There were 22 emm types identified, and the most common pharyngeal emm types were 1, 75, and 28. In contrast, the most common carriage emm types were 12, 28, and 1. The most common emm types of the secondary isolates were 4, 1, and 3.2.

DISCUSSION

This is the first prospective, population-based study of the incidence of GAS pharyngitis in a nonindigenous population in an industrialized country since the 1960s and indicates that the incidence of GAS pharyngitis in these settings is essentially unchanged in 50 years. The incidence of GAS culture–positive sore throat in children aged 5 to 12 years (13% per person-year) was very similar to the incidences of 15% and 18% per person-year for children found in 2 studies in suburban American families during the 1950s and 1960s.19,20 Higher estimates have been reported in more crowded, lower SES settings. The incidence of GAS culture–positive sore throat in children who lived in institutions during the 1950s and 1960s was 23% per person-year in the United Kingdom21 and 26% per person-year in the United States.22 In more recent studies of Maori and Pacific Island children in New Zealand and from India and Egypt, the incidence was 50%, 95%, and 30% per person-year, respectively.3,4,23

### TABLE 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>All Ages (N = 828)</th>
<th>&lt;5 y (n = 84)</th>
<th>5–12 y (n = 331)</th>
<th>13–18 y (n = 55)</th>
<th>&gt;18 y (n = 358)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sore throata</td>
<td>248</td>
<td>37 (15)</td>
<td>129 (52)</td>
<td>26 (10)</td>
<td>57 (23)</td>
</tr>
<tr>
<td>GAS culture positive</td>
<td>54 (22)</td>
<td>4 (11)</td>
<td>32 (25)</td>
<td>3 (12)</td>
<td>15 (26)</td>
</tr>
<tr>
<td>Serology confirmed</td>
<td>35 (65)</td>
<td>2 (50)</td>
<td>20 (63)</td>
<td>1 (33)</td>
<td>12 (80)</td>
</tr>
<tr>
<td>Serology negative</td>
<td>12 (22)</td>
<td>1 (25)</td>
<td>8 (25)</td>
<td>1 (33)</td>
<td>2 (13)</td>
</tr>
<tr>
<td>Serology missing</td>
<td>7 (13)</td>
<td>1 (25)</td>
<td>4 (13)</td>
<td>1 (33)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Serology unequivocalb</td>
<td>19 (35)</td>
<td>2 (50)</td>
<td>11 (34)</td>
<td>1 (33)</td>
<td>5 (33)</td>
</tr>
<tr>
<td>Secondary cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAS culture positive</td>
<td>32</td>
<td>6 (19)</td>
<td>18 (56)</td>
<td>3 (9)</td>
<td>5 (16)</td>
</tr>
<tr>
<td>Serology confirmed</td>
<td>21 (66)</td>
<td>5 (83)</td>
<td>12 (67)</td>
<td>2 (67)</td>
<td>2 (40)</td>
</tr>
<tr>
<td>Serology negative</td>
<td>3 (10)</td>
<td>0 (0)</td>
<td>2 (11)</td>
<td>0 (0)</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Serology missing</td>
<td>8 (25)</td>
<td>1 (17)</td>
<td>4 (22)</td>
<td>1 (33)</td>
<td>2 (40)</td>
</tr>
<tr>
<td>Serology unequivocalb</td>
<td>9 (28)</td>
<td>3 (75)</td>
<td>4 (22)</td>
<td>0 (0)</td>
<td>2 (40)</td>
</tr>
</tbody>
</table>

a Row percentages are given for the number of sore throat episodes, but GAS culture–positive, GAS serology values, and secondary cases are given as column percentages of the appropriate denominators.

b The unequivocal serologically proven episodes of GAS pharyngitis are a subset of the serologically confirmed episodes of GAS pharyngitis.

### TABLE 4

<table>
<thead>
<tr>
<th>Parameter</th>
<th>GAS Pharyngitis</th>
<th>Symptomatic GAS Pharyngitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>Cases, n</td>
<td>Time at Risk, Patient-Years</td>
</tr>
<tr>
<td>&lt;5</td>
<td>10</td>
<td>97</td>
</tr>
<tr>
<td>5–12</td>
<td>50</td>
<td>392</td>
</tr>
<tr>
<td>13–18</td>
<td>6</td>
<td>65</td>
</tr>
<tr>
<td>&gt;18</td>
<td>20</td>
<td>422</td>
</tr>
<tr>
<td>Overall</td>
<td>86</td>
<td>977</td>
</tr>
</tbody>
</table>

kindergarten), the hours attended, or the class size and incidence of primary cases in children (Table 5).
This is also the first study to include presence of symptoms, throat swab culture findings, and serologic confirmation to determine the relative incidences of sore throat, GAS culture–positive sore throat, and true, serologically confirmed GAS pharyngitis. All published studies used clinical criteria and culture to make the diagnosis of GAS pharyngitis except for 1,23 which used culture and serology but no clinical criteria. We found that one quarter of all people (one third of school-aged children) experienced a sore throat every year, that approximately one third of these were GAS culture positive (in all ages), and that approximately two thirds of GAS culture–positive cases were serologically confirmed GAS pharyngitis. The GAS culture–positive rate was similar to other studies in affluent populations, where rates between 24% and 36% have been reported in school-aged children.24–27 Our finding of a similarly high proportion of GAS culture–positive cases in adults suggests that, although sore throat is less common in adults than children, the chance that an adult who has a sore throat will have culture- or serologically confirmed GAS pharyngitis is similar to that of a school-aged child.

We believe that our definition of serologically confirmed GAS pharyngitis, combining unequivocal and likely cases, gives a more accurate representation of true GAS pharyngitis than unequivocal cases alone. The ideal study would include paired serology on all cases, with the first sample taken at the time of first presentation with symptoms. This was not possible in our cohort study, in which the need to minimize unnecessary blood-taking meant that we had to wait for a positive culture result before approaching participants, which was often followed by logistic delays in arranging for home visits. Therefore, the first sample was sometimes taken ≥8 days after the onset of sore throat, by which time antistreptococcal antibody titers would already have started rising in true GAS pharyngitis.13 Moreover, most children received antibiotics before the culture results, which is known to reduce the magnitude of the antibody response to GAS extracellular antigens.28–30 Our combined definition ensured that the gold standard, rising antibody titers, was used in all cases in which this could reasonably be assessed (ie, those with first samples taken within 7 days of symptom onset) and that a sensitive but less specific definition, based on population-specific ULN values (which were higher than have previously been used in this population), was used for cases in which the first specimen was delayed. Although this may have resulted in a slight overestimation of the incidence of true GAS pharyngitis, the alternative, rejecting these “likely” cases, would have resulted in a dramatic underestimation.

We found a high incidence of secondary cases, which often affected ≥1 family member. When exposed, an individual was nearly twice as likely to experience a secondary case in the family than a primary case in the community. This suggests that a single episode of GAS pharyngitis within the family has broader implications in terms of extra cost of medication and time off school and work for additional family members, which should be considered in the management of the index case.

We also found that 20% of children who were...
Younger than 5 years experienced a secondary case of GAS pharyngitis. Higher rates of GAS pharyngitis are becoming increasingly common in younger children, presumably since the advent of child care. A study from Israel surveyed 866 children who were younger than 2 years in 1999, 371 of whom had symptoms, and 98 (26%) of these were positive for GAS.35 In a child care study in the United States in 1989, 18 (27%) of 66 children who were aged <3.5 years and had symptoms were found to be GAS positive.36

The incidence rates of sore throat and GAS pharyngitis did not seem to be influenced by SES or household crowding; however, like most other studies, there was a lack of extremes of these indicators in our study population, making it difficult to examine them as independent risk factors. A lack of association with economic indicators was found in predominantly middle class American families in the 1960s19,25,33,34 and in homogeneously poor families in India.4 The much higher incidence of GAS pharyngitis in children in India4 and from Maori or Pacific Island families in New Zealand3 than was found in our study and others in affluent populations suggests that there may be an association between low SES and higher GAS infection rates, although ethnic differences may also have an influence. Most cohort studies were conducted in small to moderately sized families and were unable to demonstrate an association with crowding,4,35,36 probably because the level of crowding present in military studies that demonstrated an effect of close proximity of beds on acquisitions of GAS37 was of a greater magnitude than has been found in family studies.

CONCLUSIONS

This study provides the first reliable, population-based data on the burden of sore throat and GAS pharyngitis in an affluent population for several decades. We have confirmed that GAS pharyngitis remains as common now in school-aged children as it was 50 years ago. Our study also highlights the higher-than-expected proportion of culture-positive GAS pharyngitis in adults and the potentially important role of intrafamilial transmission. We also propose that future cohort studies use our combined definition of serologically proven GAS pharyngitis when there are substantial numbers of cases for which the first blood specimen is delayed.

ACKNOWLEDGMENTS

This work was funded by a 2002–2004 National Health and Medical Research Council (Australia) project grant.

We acknowledge the family medicine practitioners, Dr Michael Howson, Dr Peter Rankin, and Dr Robert Vorich, for assistance in the cohort study and Kris Jansen for assistance with statistical analysis and data processing.

REFERENCES

18. Cornfield D, Hubbard JP, Harris TN, Weaver R. Epidemiologic
studies of streptococcal infection in school children. 


27. Kaplan E, Top F, Dudding B, Wannamaker L. Diagnosis of streptococcal pharyngitis: differentiation of active infection from the carrier state in the symptomatic child. 

28. Rantz LA. Hemolytic streptococcal sore throat: antibody response following treatment with penicillin, sulfadiazine, and salicylates. 


32. Smith TD, Wilkinson V, Kaplan EL. Group A Streptococcus-associated upper respiratory tract infections in a day-care center. 


34. James WE, Badger GF, Dingle JH. A study of illness in a group of Cleveland families. XIX. The epidemiology of the acquisition of group A streptococci and of associated illnesses. 


38. Rantz LA. Hemolytic streptococcal sore throat: antibody response following treatment with penicillin, sulfadiazine, and salicylates. 


40. Kilbourne ED, Loge JP. The comparative effects of continuous and intermittent penicillin therapy on the formation of antistreptolysin in hemolytic streptococcal pharyngitis. 


42. Smith TD, Wilkinson V, Kaplan EL. Group A Streptococcus-associated upper respiratory tract infections in a day-care center. 


44. James WE, Badger GF, Dingle JH. A study of illness in a group of Cleveland families. XIX. The epidemiology of the acquisition of group A streptococci and of associated illnesses. 


48. Rantz LA. Hemolytic streptococcal sore throat: antibody response following treatment with penicillin, sulfadiazine, and salicylates.
Lipolysis and Insulin Sensitivity at Birth in Infants Who Are Large for Gestational Age

Fredrik S. E. Ahlsson, MD, Barbro Diderholm, MD, PhD, Uwe Ewald, MD, PhD, Jan Gustafsson, MD, PhD

Department of Women’s and Children’s Health, Uppsala University, University Children’s Hospital, Uppsala, Sweden

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. In addition to neonatal hypoglycemia, infants who are born large for gestational age are at risk for developing obesity, cardiovascular disease, and diabetes later in life. The aim of this study was to investigate glucose production, lipolysis, and insulin sensitivity in infants who were born large for gestational age to mothers without diabetes. The effect of glucagon administration on production of energy substrates was also investigated.

METHODS. Ten healthy term infants who were born large for gestational age to mothers without diabetes were studied 16 ± 8 hours postnatally after a 3-hour fast. Rates of glucose production and lipolysis were analyzed by gas chromatography–mass spectrometry following constant rate infusion of [6,6-2H2]glucose and [2-13C]glycerol. Insulin sensitivity was assessed by the Homeostasis Assessment Model. In 8 of the infants, the effect of an intravenous injection of 0.2 mg/kg glucagon was also analyzed.

RESULTS. Plasma glucose and glycerol averaged 3.8 ± 0.5 mmol/L and 384 ± 183 μmol/L, respectively. The glycerol production rate, reflecting lipolysis, was 12.7 ± 2.9 μmol/kg per min. Mean rate of glucose production was 30.2 ± 4.6 μmol/kg per min. Homeostasis Assessment Model insulin sensitivity corresponded to 82% ± 19%, β-cell function to 221% ± 73%, and insulin resistance to 1.3 ± 0.3. After glucagon administration, rate of glucose production increased by 13.3 ± 8.3 μmol/kg per min and blood glucose by 1.4 ± 0.5 mmol/L. Glycerol production decreased from 12.8 ± 3.0 to 10.7 ± 2.9 μmol/kg per min. Mean insulin concentration increased from 10.9 ± 3.0 to 30.9 ± 10.3 mU/L. There was a strong inverse correlation between the decrease in lipolysis and increase in insulin after glucagon administration.

CONCLUSIONS. Infants who are born large for gestational age show increased lipolysis and a propensity for decreased insulin sensitivity already at birth. The simultaneous increase in plasma insulin correlated strongly with the noted decrease in lipolysis, indicating an antilipolytic effect of insulin in these infants.
Obesity is a major health problem in the Western world. There has been a marked increase in the prevalence of overweight and obesity among children during the past few decades. Correspondingly, the relative number of infants who are born large for gestational age (LGA) has also increased. Data show that infants who are born LGA have a high prevalence of overweight when they reach adolescence and are at increased risk for developing cardiovascular disease and type 1 as well as type 2 diabetes in adult life. Infants who are born LGA, irrespective of its cause, are also at risk for having neonatal hypoglycemia.

At birth, the continuous placental flow of nutrients, mostly glucose and amino acids, is terminated. Before breastfeeding is established, the newborn infant has to produce glucose, mainly to meet the needs of the central nervous system. Glucose is the most important energy substrate for the brain, and during rest, the central nervous system consumes the major part of the hepatic glucose production.

Neonatal energy substrate production has been extensively studied in both infants who are appropriate for gestational age (AGA) and infants who belong to risk groups (eg, those born preterm or small for gestational age [SGA], infants of mothers with diabetes). The immediate postnatal glucose production and lipolysis are under hormonal regulation. Hepatic glucose production is stimulated by a decreased insulin/glucagon ratio, and lipolytic hydrolysis of depot fat is enhanced by the marked increase in thyrotropin that occurs during the first day of life. In contrast to the situation later in life, it is not established whether insulin has a role in the regulation of lipolysis in newborn infants.

The relation between fetal/neonatal nutrition and adult metabolic disease has been discussed extensively in recent years. It has been reported, for instance, that infants who are born SGA have increased insulin sensitivity at birth, although they may develop insulin resistance already in childhood.

Besides the risk for neonatal hypoglycemia infants who are born LGA are at risk for developing obesity and metabolic disease later in life. No information is available on neonatal insulin sensitivity or formation of energy substrates in infants who are born LGA. The aim of this study was to estimate the rates of glucose production and lipolysis and also to assess the insulin sensitivity in infants who are born LGA, irrespective of its cause, at risk for having neonatal hypoglycemia.

The Human Ethics Committee of the Medical Faculty, University of Uppsala, approved the study.

METHODS
The Human Ethics Committee of the Medical Faculty, University of Uppsala, approved the study.

## Study Infants
Ten healthy newborn term infants who had a mean gestational age of 40 ± 1.6 weeks and a mean birth weight of 4734 ± 487 g and were born LGA (4 girls) to mothers who did not have diabetes (Table 1) were recruited at the maternal ward of Uppsala University Hospital. Oral consent was obtained from both parents after oral and written information. Five of the infants were delivered by cesarean section and 5 vaginally. The pregestational BMI of the mothers averaged 29.5 ± 7 kg/m². Eight of the mothers were healthy, and 1 was receiving medication for depression (Sertraline 150 mg/day) and 1 had mild asthma treated intermittently with beclomethasone. For screening of the pregnant population, fasting blood glucose levels are measured 4 times during pregnancy (weeks 10–14, 20–24, 28–32, and 32–36). If the blood glucose exceeds 8.0 mmol/L, then an oral glucose tolerance test is performed. This was necessary for 1 woman, and the results of here oral glucose tolerance test were normal.

LGA was defined as a birth weight >2 SD for gestational age according to the Swedish fetal growth chart. Five infants were also tall for gestational age. Gestational age was determined by ultrasound examination in weeks 16 to 18 of pregnancy. The infants were studied at a postnatal age of 16 ± 8 hours. Before the study, all infants were breastfed. Two infants also received formula. The interval between the last feed and the study was at least 3 hours. No infant developed hypoglycemia during the investigation.

### Isotope Tracers
[6,6-2H₂]Glucose (isotopic purity: 98%) and [2-13C]glycerol (isotopic purity: 98%) were purchased from Cambridge Isotope Laboratories (Woburn, MA). The [6,6-2H₂]glucose and [2-13C]glycerol were dissolved in 0.9% saline in concentrations of 4.5 and 1.2 mg/mL, respectively. The tracers were tested for sterility and pyrogenicity as described previously.

### Table 1 Characteristics of the Infants

<table>
<thead>
<tr>
<th>Infant</th>
<th>Birth Weight, g</th>
<th>Birth Length, cm</th>
<th>Gestational Age, wk</th>
<th>Postnatal Age, h</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4960</td>
<td>56.5</td>
<td>42.3</td>
<td>24.0</td>
</tr>
<tr>
<td>2</td>
<td>5020</td>
<td>56.0</td>
<td>39.3</td>
<td>28.0</td>
</tr>
<tr>
<td>3</td>
<td>5390</td>
<td>57.0</td>
<td>41.4</td>
<td>19.0</td>
</tr>
<tr>
<td>4</td>
<td>4180</td>
<td>50.0</td>
<td>38.9</td>
<td>7.0</td>
</tr>
<tr>
<td>5</td>
<td>5225</td>
<td>55.0</td>
<td>41.3</td>
<td>25.0</td>
</tr>
<tr>
<td>6</td>
<td>5030</td>
<td>56.0</td>
<td>41.1</td>
<td>19.0</td>
</tr>
<tr>
<td>7</td>
<td>4200</td>
<td>53.0</td>
<td>38.0</td>
<td>10.0</td>
</tr>
<tr>
<td>8</td>
<td>4930</td>
<td>56.0</td>
<td>41.1</td>
<td>9.0</td>
</tr>
<tr>
<td>9</td>
<td>4350</td>
<td>50.0</td>
<td>38.7</td>
<td>9.0</td>
</tr>
<tr>
<td>10</td>
<td>4050</td>
<td>51.0</td>
<td>38.0</td>
<td>16.0</td>
</tr>
<tr>
<td>Mean</td>
<td>4734</td>
<td>55.0</td>
<td>40.0</td>
<td>16.4</td>
</tr>
<tr>
<td>SD</td>
<td>487</td>
<td>2.8</td>
<td>1.6</td>
<td>7.8</td>
</tr>
</tbody>
</table>

PEDIATRICS Volume 120, Number 5, November 2007 959
Study Design
The study was performed at the neonatal care unit of the University Children’s Hospital (Uppsala, Sweden). Two peripheral vein catheters were inserted, 1 for infusion of the tracers and the other for collection of blood samples. The tracers were infused for 140 minutes. Blood samples were obtained before the start of the tracer infusion and then every 10 minutes between 60 and 140 minutes (a total of 8–9 samples; 1 mL per sample corresponded to ~2.2% of the estimated blood volume). The effect of an intravenous injection of 0.2 mg/kg glucagon (Glucagon; Novo Nordisk, Bagsvaerd, Denmark; 1.0 mg/mL), given 90 minutes after the start of isotope infusion, was analyzed in 8 of the infants.

Chemical Procedures
Blood glucose was measured directly by the glucose oxidase method (ABL 735; Radiometer, Copenhagen, Denmark). The blood from the EDTA tubes was instantly centrifuged, and the plasma was frozen at −70°C until further analyzed. For measurement of plasma glycerol, an internal standard of [1,1,2,3,3-2H5]glycerol (isotopic purity: 98%), purchased from Cambridge Isotope Laboratories, was added to the plasma samples. The plasma proteins were precipitated with acetone, and after evaporation to dryness, equal amounts of pyridine and acetic anhydride were added for the preparation of the pentacetate derivative of glucose and triacetate derivative of glycerol. The isotopic enrichments of [6,6-2H2]glucose, [2-13C]glycerol, and [1,1,2,3,3-2H5]glycerol were determined by gas chromatography–mass spectrometry. A Finnigan SSQ 70 mass spectrometer (Finnigan MAT, San Jose, CA) equipped with a an HP 5890 gas chromatograph (Hewlett-Packard, Palo Alto, CA) with a nonpolar (DB1) capillary column (15 m × 0.25 mm) was used. The temperatures in the oven were changed according to a program, from 180 to 250°C and from 100°C to 140°C for glucose and glycerol, respectively. Methane was used for chemical ionization with selective monitoring of ions. The ions monitored had m/z (mass-to-charge ratio) of 331, 332, and 333, corresponding to unlabeled, 13C-labeled (M + 1), and dideuterated glucose (M + 2). For glycerol, the ions m/z 159, 160, and 164 were monitored, reflecting unlabeled glycerol, 13C-labeled glycerol (M + 1), and the 5-deuterated internal standard (M + 5). The contribution of 13C2-glucose to M + 2 was analyzed for 2 of the infants (4 and 5). This was done by gas chromatography–mass spectrometry of the saccharic acid tetraacetate derivative of glucose with monitoring of ions 347 (M) and 349 (M + 2). 13C2-glucose was shown to contribute <10% to the M + 2 enrichment of plasma glucose in both cases.

Insulin, glucagon, insulin-like growth factor 1 (IGF-1), and insulin-like growth factor–binding protein 1 (IGFBP-1) were measured in pooled samples, obtained during periods of approximate steady state. The radioimmunoassay technique was used to measure insulin, IGF-1, IGFBP-1, and glucagon (kit RB 310; Euro-Diagnostica AB, Medeon, Malmö, Sweden).

Calculations
The concentrations of plasma glycerol were calculated during periods of approximate steady state before and after injection of glucagon (mean coefficients of variation [CVs]: 8% and 10%, respectively) from the ion current ratio 159:164 by the use of a standard curve. For preparation of a standard solution, an amount of the internal standard equal to that added to the plasma samples was added to increasing amounts of unlabeled glycerol. The mean CVs for plasma glucose concentration during approximate steady state before and after administration of glucagon were 7% and 5%, respectively. Isotopic enrichments of [6,6-2H2]glucose and [2-13C]glycerol were used to calculate appearance rates of glucose and glycerol during periods of approximate steady state before and after injection of glucagon. The CVs were 4% and 2%, respectively, for glucose (m/z: 333:331) and 6% and 4%, respectively, for glycerol (m/z: 160:159). The standard curves used were prepared by gradually increasing the amounts of labeled glucose and glycerol in relation to the corresponding unlabeled compounds. The glucose production rate (GPR) and the rate of glycerol production were calculated as follows: Production rate = (i × 100/IE), where i is the infusion rate of the tracer and IE is the isotopic enrichment of the tracer in plasma (given as molar ratio [ie, labeled (tracer)/unlabeled substrate in %]). The fraction of glycerol converted to glucose and the fraction of glucose derived from glycerol were calculated from 13C enrichment of glucose reflected by an m/z of 332:331 before and after glucagon injection (mean CVs were 2% and 1.5%, respectively, during approximate steady state) as described by Patel and Kalhan. As stated already, the contribution of 13C2-glucose to M + 2 of plasma glucose was calculated in 2 of the infants as described by Hellerstein et al. Insulin sensitivity was assessed by using homeostasis model assessment (HOMA). The HOMA Calculator 2.2 program (Diabetes Research Laboratory, Oxford, United Kingdom) was used. The HOMA index correlates well with more complex measures of insulin resistance in adults. Insulin sensitivity was also assessed by calculating the plasma glucose (mg/dL)/insulin (mU/L) ratio. This ratio and the HOMA index were used by Bazaes et al for calculating insulin sensitivity in infants who were born AGA and SGA. Enteral contributions to the rates of appearance of glucose and glycerol could not be calculated, but because the mean duration of fasting was at least 3 hours, these contributions should have been minimal.
Statistical Analysis
The results are presented as mean ± SD or median and range when not normally distributed. Correlation analyses were performed with Pearson’s correlation 2-tailed test. Comparisons between measurements before and after glucagon injection were made with the paired student’s t test. Differences and correlations were considered significant at P < .05.

RESULTS
Glucose
The mean plasma glucose concentration in the infants who were born LGA was 3.8 ± 0.5 mmol/L, and the mean GPR was 30.2 ± 4.6 μmol/kg per min.

Lipolysis
The mean plasma concentration of glycerol was 384 ± 183 μmol/L, and the mean rate of glycerol production was 12.7 ± 2.9 μmol/kg per min. The fraction of glycerol converted to glucose averaged 59% ± 20%, which corresponded to 13% ± 5% of the total glucose production (Table 2).

Hormones
Serum concentrations of insulin, glucagon, IGF-1, and IGFBP-1 were 10.8 ± 2.8 mU/L, 52 pmol/L (range: 34–107 pmol/L), 38 ± 17 μg/L, and 231 ± 79 μg/L, respectively.

Insulin Sensitivity
The glucose/insulin ratio was 6.6 ± 1.6. As calculated according to the HOMA, insulin sensitivity was 82% ± 19%, β-cell function was 221% ± 73%, and insulin resistance was 1.3 ± 0.3.

Effects of Glucagon
After injection of glucagon, the mean GPR increased by 13.3 ± 8.3 μmol/kg per min (P < .05; Fig 1), and the mean blood glucose level increased by 1.4 ± 0.5 mmol/L (P < .05; ie, by 44% and 37%, respectively). Simultaneously, the mean rate of glycerol production decreased by 16% from 12.8 ± 3.2 to 10.7 ± 2.9 μmol/min per kg (P < .05; Fig 2). There was also a decrease in the proportion of glucose generated from glycerol (P < .05). The mean insulin concentration increased from 10.9 ± 3.0 to 30.9 ± 10.3 mU/L.

Correlations
Both the rate of lipolysis (r = 0.680; P < .05) and GPR (r = 0.716; P < .05) correlated with birth weight. In addition, GPR correlated with the blood glucose level (r = 0.680; P < .05). There was a strong inverse correlation between the decrease in lipolysis and the increase in insulin found after administration of glucagon (r = −0.808; P = .015; Fig 3). Furthermore, the level of IGF-1 correlated inversely with postnatal age (r = −0.749; P < .05). No correlations were found between markers for insulin resistance (HOMA index or glucose/insulin ratio) and auxologic parameters (birth weight, BMI, or ponderal index).

DISCUSSION
Much interest is being focused today on the developmental origin of adult disease, particularly on the long-term metabolic consequences of being born SGA. Data have also shown that being born LGA may predispose to overweight, diabetes, and cardiovascular disease in adult life. The relative number of infants who are born LGA is increasing, and one of the underlying reasons for this increase is overweight and obesity among pregnant women. The mothers in this study had a mean prepregnancy BMI close to that associated with obesity. Only limited information is available on neonatal metabolism in infants who are born LGA. In this study, energy substrate production and insulin sensitivity were investigated during the first day of life in a group of healthy term infants who were born LGA. In addition, the effect of glucagon administration on energy substrate produc-

### TABLE 2
Plasma Glucose, GPR, Plasma Glycerol, Rate of Glycerol Production, Fraction of Glycerol Converted to Glucose, and Fraction of Glucose Derived From Glycerol in Newborn Infants Who Were Born LGA (n = 10)

<table>
<thead>
<tr>
<th>Infant</th>
<th>P-glucose, mmol/L</th>
<th>GPR, μmol/kg per min</th>
<th>P-glycerol, μmol/L</th>
<th>Glycerol Production Rate, μmol/kg per min</th>
<th>Glycerol to Glucose, %</th>
<th>Glucose From Glycerol, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.2</td>
<td>33.9</td>
<td>424</td>
<td>13.4</td>
<td>73</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>3.4</td>
<td>33.6</td>
<td>296</td>
<td>16.4</td>
<td>29</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>4.6</td>
<td>35.6</td>
<td>829</td>
<td>15.5</td>
<td>52</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>3.4</td>
<td>27.4</td>
<td>381</td>
<td>8.5</td>
<td>72</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>3.9</td>
<td>36.6</td>
<td>397</td>
<td>13.2</td>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>3.4</td>
<td>24.7</td>
<td>512</td>
<td>16.5</td>
<td>55</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>3.1</td>
<td>23.2</td>
<td>269</td>
<td>11.5</td>
<td>77</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>4.1</td>
<td>30.2</td>
<td>232</td>
<td>10.6</td>
<td>43</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>3.6</td>
<td>28.9</td>
<td>309</td>
<td>8.5</td>
<td>85</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>3.8</td>
<td>28.0</td>
<td>193</td>
<td>12.5</td>
<td>75</td>
<td>19</td>
</tr>
<tr>
<td>Mean</td>
<td>3.8</td>
<td>30.2</td>
<td>384</td>
<td>12.7</td>
<td>59</td>
<td>13</td>
</tr>
<tr>
<td>SD</td>
<td>0.5</td>
<td>4.6</td>
<td>183</td>
<td>2.9</td>
<td>20</td>
<td>5</td>
</tr>
</tbody>
</table>
tion was studied. It would have been desirable to compare the data with those of a matched control group of term infants who were AGA, but owing to the problem of recruiting healthy newborn infants for metabolic studies, this was not possible. The results are therefore compared with literature data and also with previous data from our own laboratory on infants who were AGA.

**Lipolysis**

Lipolysis is an important source of energy in the newborn infant. One principal finding in this study was that infants who were born LGA had an efficient lipolysis. As compared with infants who were born AGA, studied previously in our laboratory at a postnatal age of 4 hours, the infants who were born LGA had an almost 50% higher rate of lipolysis, as reflected by the glycerol production. The same was found when the data were compared with those of Patel and Kalhan, who studied at a postnatal age of 21 hours (close to that of the infants in this study) term infants who were born AGA. The infants studied by Patel and Kalhan had fasted longer than our infants who were born LGA, which further strengthens that there is a difference in rate of lipolysis between infants who are born LGA and AGA. The difference becomes even more pronounced when our data are compared with those of Bougnères et al, who studied infants at a postnatal age of 22 hours who were born AGA. Data on the body composition of infants who were born LGA show that their proportion of body fat is increased compared with infants who were born AGA and SGA. In this study, there was a strong correlation between the rate of lipolysis and birth weight. This is in

![Figure 1](image1.png)

**FIGURE 1**
Effect of glucagon on GPR ($n = 8; P < .05$).

![Figure 2](image2.png)

**FIGURE 2**
Effect of glucagon on rate of glycerol production (lipolysis) ($n = 8; P < .05$).
agreement both with the finding in adults, showing an association between lipolysis and body weight, and with recent results from our laboratory on infants who were born SGA, indicating that lipolysis depends on the amount of stored fat. Reduced adipose tissue insulin sensitivity could be another mechanism contributing to the increase in lipolysis. Insulin is a widely known inhibitor of lipolysis in adults, but this role has been questioned in the newborn. Previous data showed unimpaired lipolysis despite increased insulin concentrations in infants of mothers with diabetes, findings that contrast to the antilipolytic effects exerted by corresponding levels of insulin in adults. The lack of an antilipolytic effect of insulin in newborns has been interpreted as a protective mechanism for securing their supply of energy substrates during hyperinsulinemia; however, in contrast to earlier reports, these data indicate that insulin does in fact have some regulatory role with regard to lipolysis in newborns, at least at higher levels, because a fairly strong correlation was noted between the increase in insulin and the decrease in lipolysis after glucagon administration.

**Glucose**

The mean GPR in our study cohort was at the high end of the range reported for infants who are born AGA, 28,33,34 The correlation between birth weight and length shows that the infants were symmetrically large. Consequently, the high GPR should reflect brain size. This is supported by the finding of a mean head circumference corresponding to +1.5 SD. Glycogenolysis is probably still an important source of glucose 16 hours post partum in infants who are born AGA, because the proportion of glycerol converted to glucose accounted for only ~10% of the total GPR. The contribution made by glycerol to glucose production was smaller than that reported by Patel and Kalhan for infants who were born AGA and were of the same postnatal age but close to that found by our group in infants who were born AGA and had a postnatal age of 4 hours.

**Insulin Sensitivity**

Previous studies indicated that infants who are born LGA are at higher risk for developing both type 1 and type 2 diabetes than are infants who were born AGA. In our infants, the mean level of insulin was higher than that reported previously for term infants who were born AGA13,18,28,36,37 This and the fact that -cell function according to HOMA was increased compared with that reported by Bazaes et al for infants who were born AGA indicate a decreased insulin sensitivity at birth in infants who are born LGA, analogous to that in older children with overweight or obesity.

The risk for type 1 diabetes in infants who are born LGA has been suggested to be attributable to a high -cell activity, in turn leading to increased expression of antigens that are associated with diabetes. This study demonstrates that infants who are born LGA in fact have a high -cell activity. A decreased insulin sensitivity already at birth in combination with a propensity for obesity later in life may predispose for type 2 diabetes in this particular group of infants.

That GPR showed no correlation to the insulin level or the insulin/glucacon ratio is in agreement with data reported by Hawdon et al indicating that the neonatal hepatocyte may be insensitive to insulin. The strong
increase in insulin after glucagon administration indicates a good pancreatic secretory response, a finding that contrasts to the report in the literature\textsuperscript{19} that only a weak insulin response to the increase in glucose was obtained after glucagon treatment of hypoglycemia in infants who were born AGA.

**IGF-1**

It is widely known that IGF-1 is important for fetal growth. Numerous studies have shown that the IGF-1 level in cord blood\textsuperscript{40-42} correlates with birth weight, but data on IGF-1 levels during the first 24 to 48 hours of life are limited; however, some data on IGF-1 levels during the first day of life are available. For example, our infants who were born LGA had markedly higher IGF-1 levels than those found in infants who were born AGA at a postnatal age of 24 hours by de Zegher et al\textsuperscript{43} and Giudice et al.\textsuperscript{44} Furthermore, the inverse correlation between IGF-1 level and postnatal age observed in our study is in agreement with the reported decline in IGF-1 during the first days of life.\textsuperscript{43,44}

It has been reported that the IGFBP-1 concentration is decreased in infants who are born LGA.\textsuperscript{44} The infants who were born LGA in this investigation had lower IGFBP-1 levels than newborn infants who were born SGA and recently studied in our laboratory, probably reflecting the effect of a higher insulin tone on the liver in the infants who were born LGA.

**Effect of Glucagon**

Infants who are born LGA both to mothers with and without diabetes are prone to develop neonatal hypoglycemia.\textsuperscript{10,45} This occurs despite increased stores of fat\textsuperscript{20} and liver glycogen\textsuperscript{46}; therefore, administration of glucagon should be a suitable treatment for hypoglycemia in these infants. After intravenous injection of glucagon, the blood glucose level and endogenous glucose production increased in all infants studied. Data on the conversion of glycerol to glucose indicate that under these conditions, glycogenolysis is stimulated more than gluconeogenesis. This is in agreement with recent findings by van Kempen et al,\textsuperscript{47} who studied the effect of glucagon in moderately preterm (30 weeks) infants who were born AGA and SGA; however, if glucagon is used for treatment of neonatal hypoglycemia, then the risk for nausea in the treated infants has to be considered.

**CONCLUSIONS**

Infants who were born LGA had an increased lipolysis and a propensity for decreased insulin sensitivity already at birth as compared with infants who were born AGA. Administration of glucagon markedly increased the blood glucose levels and glucose production. The simultaneous increase in insulin correlated strongly with the decrease in lipolysis after glucagon injection, indicating an antilipolytic effect of insulin in newborn infants who were born LGA.

**ACKNOWLEDGMENTS**

This study was supported by grants from the Medical Research Council, the Gillberg’s Foundation and, Wera Ekström Foundation.

We are grateful to Elisabeth Söderberg for excellent technical assistance. We also thank Cecilia Ewald and the staff of the NICU, Uppsala University Children’s Hospital, for skillful assistance and Yvonne Strömberg, Elvi Sandberg, Kerstin Brismar, and Claes-Göran Östensson (Department of Molecular Medicine, Karolinska Institute, Stockholm, Sweden) for insulin, glucagon, IGF-1, and IGFBP-1 analyses.

**REFERENCES**

7. Wei JN, Sung FC, Li CY, et al. Low birth weight and high birth weight infants are both at an increased risk to have type 2 diabetes among schoolchildren in Taiwan. Diabetes Care. 2003; 26:343–348
43. de Zegher F, Devlieger H, Veldhuis JD. Properties of growth hormone and prolactin hypersecretion by the human infant on the day of birth. J Clin Endocrinol Metab. 1993;76:1177–1181
Current Definitions of Hypotension Do Not Predict Abnormal Cranial Ultrasound Findings in Preterm Infants

Catherine Limperopoulos, PhD\textsuperscript{a,b}, Haim Bassan, MD\textsuperscript{a}, Leslie A. Kalish, ScD\textsuperscript{c}, Steven A. Ringer, MD\textsuperscript{d}, Eric C. Eichenwald, MD\textsuperscript{d}, Gene Walter, REEGT\textsuperscript{e}, Marianne Moore, BA, RN\textsuperscript{a}, Matthew Vanasse, RRT\textsuperscript{a}, Donald N. DiSalvo, MD\textsuperscript{e}, Janet S. Soul, MD, CM\textsuperscript{a}, Joseph J. Volpe, MD\textsuperscript{a}, Adr\`e J. du Plessis, MBChB, MPH\textsuperscript{a}

\textsuperscript{a}Fetal-Neonatal Neurology Research Group, Department of Neurology, and \textsuperscript{b}Clinical Research Program and Department of Pediatrics, Children’s Hospital Boston and Harvard Medical School, Boston, Massachusetts; \textsuperscript{c}Department of Neurology and Neurosurgery and School of Physical and Occupational Therapy, McGill University, Montreal, Quebec, Canada; Departments of \textsuperscript{d}Newborn Medicine and \textsuperscript{e}Radiology, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Hypotension is a commonly treated complication of prematurity, although definitions and management guidelines vary widely. Our goal was to examine the relationship between current definitions of hypotension and early abnormal cranial ultrasound findings.

METHODS. We prospectively measured mean arterial pressure in 84 infants who were $\leq 30$ weeks’ gestational age and had umbilical arterial catheters in the first 3 days of life. Sequential 5-minute epochs of continuous mean arterial pressure recordings were assigned a mean value and a coefficient of variation. We applied to our data 3 definitions of hypotension in current clinical use and derived a hypotensive index for each definition. We examined the association between these definitions of hypotension and abnormal cranial ultrasound findings between days 5 and 10. In addition, we evaluated the effect of illness severity (Score for Neonatal Acute Physiology II) on cranial ultrasound findings.

RESULTS. Acquired lesions as shown on cranial ultrasound, present in 34 (40%) infants, were not predicted by any of the standard definitions of hypotension or by mean arterial pressure variability. With hypotension defined as mean arterial pressure $< 10$th percentile ($< 33$ mm Hg) for our overall cohort, mean value for mean arterial pressure and hypotensive index predicted abnormal ultrasound findings but only in infants who were $\geq 27$ weeks’ gestational age and those with lower illness severity scores.

CONCLUSIONS. Hypotension as diagnosed by currently applied thresholds for preterm infants is not associated with brain injury on early cranial ultrasounds. Blood pressure management directed at these population-based thresholds alone may not prevent brain injury in this vulnerable population.
Hypotension is commonly diagnosed in preterm infants during the first days of life. However, there is great inconsistency in the diagnosis and management of hypotension in preterm infants, largely because of the lack of a reliable and consistently applied definition of hypotension. Disparities in the diagnosis of hypotension in preterm infants are in large part attributable to incomplete understanding of the more complex and prolonged circulatory transition from fetal to stable neonatal hemodynamics. Coincident with this period of systemic circulation, peripheral vascular resistance of the extrauterine circulatory pathways (eg, ductus arteriosus) may be more delayed in preterm infants. During this transition, the immature myocardium confronts an abrupt increase in afterload as the low-resistance placental bed is replaced by the high peripheral vascular resistance of the extraterine circulation. In addition, closure of fetal circulatory pathways (eg, ductus arteriosus) may be more delayed in preterm infants. Coincident with this period of systemic circulatory adaptation is a period of particular risk for brain injury in preterm infants. Consequently, it is not surprising that hemodynamic factors have frequently been associated with the development of early brain injury in the preterm infant.

Normal blood pressure (BP) provides appropriate perfusion pressure in end organs to maintain their functional and structural integrity. Hypotension refers to BP levels that are too low to achieve this goal. Although in theory hypotension is an organ-specific diagnosis, the particular vulnerability of the immature brain to hypoperfusion has made brain injury in the preterm infant a leading end point in the pursuit of a reliable clinical definition of hypotension; however, for a number of potential reasons, such a goal has been elusive, and several different definitions of hypotension are in clinical use. In fact, the role of hypotension in prematurity-related brain injury remains a contentious issue, as evidenced by recent calls for a radical reassessment of the diagnosis and management of hypotension in the preterm infant. To evaluate the association between hypotension and brain injury among critically ill preterm infants, we tested the ability of various definitions of hypotension that are in clinical use to predict early-life ultrasound evidence of brain injury.

Methods
As part of a previous study that investigated cerebral pressure autoregulation in preterm infants, we prospectively enrolled a cohort of preterm infants who were ≤30 weeks’ gestational age (GA) and whose hemodynamic instability required an umbilical arterial catheter for continuous BP monitoring. Because we wished to confine our focus to the ultrasound correlates of hemodynamic disturbances during the transitional circulation, we included infants with periods of continuous BP recording during the first 3 days of life. We were interested in the association between actual BP and brain injury and therefore included infants regardless of whether they were being treated for hypotension. We included only infants with at least 1 cranial ultrasound from day 5 to day 10. To exclude the effects of preceding antenatal insults, we excluded infants with cranial ultrasound evidence of long-standing injury (eg, cystic parenchymal lesions). Likewise, to avoid the cumulative effects of later insults (eg, apnea and bradycardia, infection), we confined our ultrasound outcomes to the first 5 to 10 days. We recognize that nonhemorrhagic brain injuries (particularly diffuse white matter injury) may be delayed in their ultrasound appearance and that the effect of less severe insults may be missed by this approach. We also excluded infants with known or suspected cerebral dysgenesis, obvious dysmorphic syndromes, or a known chromosomal disorder. We obtained written informed consent for all patients. The institutional review board of the Brigham and Women’s Hospital approved the study.

We categorized our population into 2 GA groups: 23 to 26 weeks and 27 to 30 weeks. We confined our BP analyses to measurements during the first 3 days of life, defined as the first 3 consecutive 24-hour intervals after birth. This restriction was made because the use of intraarterial monitoring decreases significantly thereafter and because the vast majority of germinal matrix–intraventricular hemorrhage (GM-IVH) occurs during this time. The BP transducer was maintained at a consistent midthoracic level. We excluded all BP data that were recorded during periods of line clamping (eg, arterial blood gas sampling), as well as during movement of the BP transducer.

BP Measurements
We measured continuous mean arterial BP (MAP) at 2 Hz for up to 12 hours on each of the first 3 days. We divided these continuous MAP data into sequential 5-minute epochs. The mean MAP was calculated for each epoch, as was the coefficient of variation (CV) of all MAP measurements during each epoch. The effects of MAP over the longer term were measured using 2 approaches. First, we summarized these short-term measures of MAP level and MAP variability over longer time periods by deriving the hypotensive index (HOI) and coefficient of variation index (CVI), described later. In addition, a mean MAP was calculated for each of the first 3 days of life by averaging all MAPs recorded during each day; then by averaging the 3 daily mean MAPs, we calculated an overall mean MAP for the 3-day period for each patient.

We then tested 3 definitions of hypotension that are used in clinical practice for their ability to predict early abnormal ultrasound findings. These definitions were MAP of (1) <30 mm Hg, (2) less than the infant’s GA in weeks, and (3) <10th percentile of MAP for birth weight and postnatal age on published normative data. For reasons discussed in “Results,” we also considered a fourth definition of hypotension: MAP < 10th...
percentile (33 mm Hg) in our overall study population during the first 3 days of life.

We derived the HOI in the following manner. First, consecutive 5-minute epochs from the time of birth were formed for each infant. We excluded epochs that did not have assessable MAPs recorded for at least 4 of the 5 possible minutes. By applying each of the definitions of hypotension to the mean MAP during the 5-minute epoch, each epoch was categorized as either a hypotensive epoch or a nonhypotensive epoch. The HOI was then derived as the proportion of all assessable 5-minute epochs that were hypotensive. Applying this algorithm to the 4 definitions of hypotension yielded 4 HOIs for each infant. We then compared 3-day HOI as well as the 3-day mean MAP between infants with normal versus abnormal cranial ultrasound findings, as defined later. We also calculated HOI during the first 12 hours of life to investigate the possible importance of BP during the very early newborn period.

Similarly, we examined the association between BP variability and abnormal ultrasound findings as follows. First, we calculated the CV of all MAP readings within each epoch. Epochs with a CV of >8% were categorized as having high variability, and a CVI was calculated as the proportion of all assessable epochs with high variability. We also conducted a sensitivity analysis by calculating CVIs after redefining high variability as CV > 6% and CV > 10%. We chose these cut points because they were integer values at high percentiles of all CVs, across all epochs in all infants. The cut points (6%, 8%, and 10%) corresponded to the 80th, 93rd and 98th percentiles, respectively. The CVIs were compared between infants with normal versus abnormal cranial ultrasound findings.

Finally, we compared abnormal cranial ultrasound findings between preterm infants who were treated for hypotension and those who were not treated. The BP thresholds that are used to guide pressor use in our NICU are GA and postnatal age dependent. Specifically, during the first 48 hours after birth, infants who are <30 weeks’ GA are treated for hypotension when MAP is persistently less than GA (in weeks). Hereafter, treatment is started when the MAP is persistently <2 to 4 mm Hg above GA. Inevitably, there are variations in practice, because clinicians take into account not only BP but also other clinical features. The initial treatment of hypotension is usually intravenous fluid (volume expanders) in 1 to 2 boluses of 10 mL/kg each. For persistent hypotension, the first-line pressor medication is dopamine, up to an infusion rate of 30 μg/kg per min. Thereafter, dobutamine, then epinephrine, and finally corticosteroids are used, if needed. Treatment for hypotension was characterized as treatment during each of the first 3 days of life, treatment during any of the first 3 days, and the number of days of treatment. We also evaluated the association between these treatment modalities and BP.

Severity of Illness
To determine the severity of illness in our population during the early neonatal period, we applied the widely used and validated Score for Neonatal Acute Physiology II (SNAP-II) using data from the first 12 hours of life. Mean BP forms part of the SNAP-II; therefore, to evaluate illness severity independent of a BP effect, we modified the SNAP-II by calculating the mean MAP for our study population and applying this value for all infants in the BP item of SNAP-II. We used this modified, “BP-neutral” SNAP-II to evaluate the role of illness severity in the treatment of BP, as well as the association between illness severity and abnormal cranial ultrasound findings.

Cranial Ultrasound Criteria
We used as our primary outcome variable predefined features of abnormality on routine cranial ultrasound studies (Acuson Sequoia; Siemens Medical Solutions, Malvern, PA) performed on a clinically scheduled protocol. The Brigham and Women’s Hospital NICU uses a clinical protocol for the timing of ultrasound studies (day 3, day 7, and day 30 after birth) in preterm infants. However, additional studies are performed whenever clinically indicated; therefore, in sick populations such as ours, the timing of ultrasound study is seldom consistent. Because the presence of hemorrhagic and major hypoxic-ischemic injuries that occur within the first 3 days should be evident by ultrasound studies performed between 5 and 10 days, we based our outcome analysis on a single ultrasound study performed on or closest to 5 days and before 10 days after birth. The diagnosis of abnormal cranial ultrasound findings was based on a widely used grading scheme and made by 2 experienced cranial ultrasound readers (Drs DiSalvo and du Plessis) who were blinded to the clinical and BP data. We defined 3 ultrasound outcome variables: GM-IVH (any grade), parenchymal abnormalities (defined as cerebral and/or cerebellar echodensities), and an overall outcome (GM-IVH and/or parenchymal abnormalities). Echodensities were defined as areas of bright ultrasound signal, approaching or at the density of the choroid plexus and clearly distinct from the surrounding parenchyma. We considered findings to be echodensities only when they were evident on >1 ultrasound slice in the angled sagittal or coronal views. In all cases, regions of echodensity were associated with abnormalities on subsequent ultrasound and/or MRI studies.

Clinical Data Collection
We collected demographic, prenatal, intrapartum, and postnatal data on all infants in a prospective manner using standardized data forms. Demographic data included GA at birth, birth weight, and gender. Maternal data included single versus multiple gestation, pregnancy-induced hypertension, prenatal infection, maternal...
chorioamnionitis (clinical and/or pathologic diagnosis), and use of tocolysis and antenatal steroids. Clinical chorioamnionitis was entered when this diagnosis was made by the treating obstetrician. Intrapartum factors included mode of delivery, maternal fever (>38°C), vaginal bleeding. Apgar score at 5 minutes, and need for respiratory resuscitation. Early postnatal data collected for the first 3 days of life included arterial blood gases (pH, PCO2, and PO2), blood cell counts and platelets, need for cardiopulmonary support (use of volume expanders and/or pressor-inotropic agents), indomethacin for a patent ductus arteriosus, maximum infant temperature, and infection (confirmed by positive blood, urine, or mucosal culture or clinical suspicion resulting in a full course of antibiotic treatment despite negative cultures).

To evaluate laboratory results as binary covariates, we used clinically meaningful cut points to categorize each variable as follows: minimum pH < 7.2, minimum PO2 < 45 mm Hg, minimum PCO2 < 30 mm Hg, maximum PCO2 > 60 mm Hg, maximum white blood cell count > 20 000/µL, minimum hemoglobin level < 10 g/dL, and minimum platelet count < 100 000/µL.

Statistical Methods
The association between a BP index (eg, HOI) and each cranial ultrasound outcome variable was assessed by comparing the index distribution between infants with versus without the abnormal ultrasound finding using the 2-sample Wilcoxon test. To express the magnitude of these associations on a scale that is directly comparable across different indices, we also report the area under the receiver operating characteristic curve (AUROC), which is closely related to the Wilcoxon test and has the following interpretation. The AUROC is the probability that an infant with the abnormality has a higher index value than an infant without the abnormality. An AUROC of 0.50 suggests that the index has no predictive value, and an AUROC of >0.50 suggests a positive association. Because higher values of HOI or CVI are hypothesized to confer higher risk whereas lower values of BP confer higher risk, we reversed the scale when analyzing mean MAP so that AUROCs of >0.50 are always interpretable as an association in the hypothesized direction.

To compare mean MAP across days 1 through 3 and between GA strata, we fit a repeated measures model with day of life as the repeated factor and GA as a fixed factor. We used an autoregressive correlation structure. We then constructed Wald tests from linear contrasts to compare the covariate distributions in infants with cranial ultrasound outcomes compared across covariate subgroups with Fisher’s exact test. For continuous covariates, we tested associations with outcomes by comparing the covariate distributions in infants with versus without the outcome using the 2-sample Wilcoxon test.

We evaluated whether a covariate confounded the association between HOI or mean MAP and cranial ultrasound outcome as follows. We considered only covariates that had suggestive associations with outcomes as indicated by P < .10. We then compared the association evaluated in unadjusted analysis (using the Wilcoxon test) with the same association evaluated in adjusted analyses, where we adjusted for each of the covariates using a van Elteren test. The van Elteren test is a stratified version of the Wilcoxon test. All tests are 2-sided and P < .05 is considered significant.

RESULTS
Included in the analysis were 84 infants who met our entry criteria.

Baseline Characteristics
Table 1 summarizes the clinical characteristics of pregnancy, labor, and delivery. Preterm infants ranged in GA from 23 to 30 weeks and in birth weight from 460 to 1490 g. By our inclusion criteria, umbilical arterial lines were present in all infants for BP management. The mean age at onset of BP recording was 11.5 hours; however, studies started within 12 hours in 50 (60%) of 84 patients and within 24 hours of birth in all but 2. Although we aimed to record 12 hours of continuous MAP data per day during the first 3 days of life, personnel and equipment availability, as well as artifact removal and exclusion of incompletely recorded epochs, reduced the MAP data that were available for analysis. Reliable BP data were available on days 1 through 3 from 82, 78, and 60 infants for a mean (range) of 5.20

<table>
<thead>
<tr>
<th>Clinical Characteristic</th>
<th>n (%) or Median (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous preterm labora</td>
<td>63 (75)</td>
</tr>
<tr>
<td>Chorioamnionitis (N = 69)</td>
<td>14 (17)</td>
</tr>
<tr>
<td>Intrapartum or antepartum hemorrhage (N = 83)</td>
<td>74 (88)</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>26 (23–30)</td>
</tr>
<tr>
<td>GA, wk</td>
<td>880 (460–1490)</td>
</tr>
<tr>
<td>Birth weight, g</td>
<td>50 (60)</td>
</tr>
<tr>
<td>Male gender</td>
<td>57 (68)</td>
</tr>
<tr>
<td>Singleton</td>
<td>7 (0–9)</td>
</tr>
<tr>
<td>Cesarean delivery</td>
<td>55 (65)</td>
</tr>
<tr>
<td>Labor</td>
<td>17 (85)</td>
</tr>
<tr>
<td>Respiratory resuscitationb</td>
<td>6 (7)</td>
</tr>
</tbody>
</table>

N = 84 except where indicated.
a Previous preterm labor during this pregnancy.
b In the delivery room.
Abnormal Cranial Ultrasound Findings

The mean age at the target cranial ultrasound study was 7.4 days. On this ultrasound study, abnormalities were identified in 34 (40%) of 84 preterm infants. Thirty-one (37%) developed GM-IVH grades 1 through 3. Because 3 cranial ultrasound studies had suboptimal cerebellar views, 81 cranial ultrasound studies were assessable for parenchymal abnormalities, 14 (17%) of which had cerebral echodensities and/or cerebellar lesions (9 cerebral only, 4 cerebellar only, and 1 both).

Relationship of MAP With GA and Postnatal Age

Table 2 describes the relationship among mean MAP, GA, and postnatal age during each of the first 3 days and for the overall 3-day period. Generally, those in the lower GA group had lower mean MAP, although this was statistically significant only on days 2 and 3 ($P = .001$ for both). There was also a statistically significant increase in mean MAP during the first 3 days of postnatal life ($P = .01$ among infants with GA 23–27 weeks, $P < .001$ among infants with GA 27–30 weeks). However, the increase in MAP during the 3 days was much less than that previously reported,12 and we were not able to estimate daily 10th percentiles with precision adequate for creating separate daily criteria for hypotension; therefore, we used the overall 3-day data set to derive the 10th percentile of mean MAP.

Relationship Between Hypotension and Early Abnormal Cranial Ultrasound Findings

When the criterion for hypotension was based on the 3rd percentile of mean MAP, for example, 81% of infants had 0 epochs with mean MAP less than their GA. Furthermore, on the basis of the HOI derived for each of these thresholds, we found no significant associations between HOI and abnormal cranial ultrasound findings (Table 3).

Our population of infants likely differed in important ways (including overall systemic morbidity) from infants that were used to generate existing normative BP databases12,23,27; therefore, given the lack of association between hypotension defined by these criteria and abnormal cranial ultrasound findings in our infants, we considered a fourth definition of hypotension, MAP < 33 mm Hg, which is the 10th percentile of the overall 3-day mean MAP in our population, and derived the hypoten-

<table>
<thead>
<tr>
<th>GA, wk</th>
<th>AUROC Value Per Early Cranial Ultrasound Outcome ($P$)</th>
<th>IVH</th>
<th>Parenchymal</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1O1-30</td>
<td>All 0.46 0.46 0.49</td>
<td>0.54 0.50 0.54</td>
<td>0.50 0.49 0.50</td>
<td></td>
</tr>
<tr>
<td>H1O1-30</td>
<td>23–26 0.47 0.66 0.47</td>
<td>0.46 0.46 0.46</td>
<td>0.50 0.49 0.50</td>
<td></td>
</tr>
<tr>
<td>H1O1-30</td>
<td>27–30 0.47 0.66 0.47</td>
<td>0.46 0.46 0.46</td>
<td>0.50 0.49 0.50</td>
<td></td>
</tr>
<tr>
<td>H1O1-30</td>
<td>All 0.54 0.65 0.54</td>
<td>0.46 0.46 0.46</td>
<td>0.50 0.49 0.50</td>
<td></td>
</tr>
</tbody>
</table>

Only $P < .1$ is shown. AUROC values $>0.50$ suggest an association between low BP (or high BP variability) and abnormal ultrasound findings; H1O1-30 indicates threshold for hypotension is MAP < 33 mm Hg.
sive index for this threshold (HOI-33). Using this criterion, 19% of the infants had HOI-33 equal to 0 (ie, no hypotensive 5-minute epochs). We also considered the overall mean MAP for days 1 through 3 of life and the CVI as potential correlates of abnormal cranial ultrasound findings. For the remainder of our analyses, we used 33 mm Hg as the threshold of hypotension and HOI-33 as the hypotensive index.

AUROCs for the different hypotensive indices and for the overall mean MAP and CVI are shown in Table 3. The results are stratified by GA (23–26 weeks: n = 44; 27–30 weeks: n = 40). High HOI-33 and low overall mean MAP are each associated with abnormal cranial ultrasound outcomes; however, the association is confined to the GA group of ≥27 weeks. The similar AUROC values for these 2 measures of BP suggest that they perform very similarly as correlates of abnormal cranial ultrasound findings. None of the HOI using cut points based on published normative data was associated with abnormal cranial ultrasound findings. The CVI based on a CV = 8% cut point was also not associated with abnormal cranial ultrasound findings (Table 3). In a sensitivity analysis, CVIs using larger and smaller cut point values were also not predictive of abnormal cranial ultrasound findings (data not shown).

To examine features of BP confined to the earliest hours, we calculated HOI and CVI using only data from the first 12 hours after birth (n = 50). If the early BPs were associated with abnormal outcomes, then one might expect that the AUROCs based on BP indices calculated from these first hours would be larger than the AUROCs calculated from the entire 3-day period. Generally, we did not see such a pattern, and in no case was an early HOI or CVI significantly higher among infants with abnormal cranial ultrasound outcomes (Table 3). In a related analysis, we calculated the daily mean MAP separately in infants with versus without abnormal ultrasound findings (Fig 1). We did not find a larger difference on day 1, as would be expected if the earliest BPs were best able to predict outcomes.

**Effects of Volume Expander and Pressor-Inotrope Use**

For our analysis of volume expanders, we focused on dosages of >10 mL/kg per day, which were administered to 63 (75%) of our infants, whereas 60 (71%) infants required pressor-inotrope support at some point during the 3-day study period. Dopamine was the first-line pressor-inotrope used in all infants. During the 3-day study period, 1 infant also received dobutamine (days 1 and 2), 1 received epinephrine (day 1), and 2 received hydrocortisone (1 on day 2 only and the other on days 1 and 2). Only 1 patient received 3 pressors (days 1 and 2). Given the small number of patients on multiple pressor support, additional analysis was not conducted. The numbers of infants who required volume expanders and pressor-inotropes on each study day, as well as the numbers of infants who were treated for 1, 2, or all 3 days with each of these interventions, are shown in Table 4. As shown in Table 4, infants who were treated with volume expanders or pressor-inotropes had significantly lower mean MAP values for each of the study days and for the study period overall.

Next, we considered the association between volume expanders and pressor-inotrope support, and abnormal cranial ultrasound findings. The use of volume expanders was not significantly associated with any abnormal cranial ultrasound finding on day 1 or with either GM-IVH or parenchymal echodensities on day 2; however, volume expander use on day 2 was significantly associ-
associated with overall abnormal ultrasound findings \((P = .03)\), and volume expander use on day 3 was significantly associated with all 3 abnormal ultrasound outcomes \((P < .005\) for each outcome). There was no significant association between abnormal cranial ultrasound findings and pressor-inotrope treatment (any versus no treatment) or treatment on any of the 3 study days.

**Influence of Other Clinical Variables on the Relationship Between MAP and Abnormal Cranial Ultrasound Finding**

The population selected for this study (preterm infants who required umbilical arterial BP monitoring) represents a particularly sick population, often with multiorgan dysfunction; therefore, we examined the role of selected clinical and laboratory variables in the relationship between MAP and abnormal cranial ultrasound findings. Thirty (36%) infants were treated on \(\geq 1\) days during the study period with indomethacin for patent ductus arteriosus. Neonatal infection was diagnosed in 4 (5%) infants during the study period. Arterial blood gas analyses during the first 3 days of life, shown as median (range), were minimum pH 7.2 (6.8–7.4), minimum Po2 (mm Hg) 42.6 (23.3–63.5), minimum Pco2 (mm Hg) 31.9 (17.2–47.9), and maximum Pco2 (mm Hg) 61.6 (36.5–169.5). The median (range) of the maximum white blood count \((10^9/\mu L)\) was 11.3 (3.3–72.0) and of minimum platelets \((10^9/\mu L)\) was 168 (59–439). We selected only clinical variables with \(P < .10\) when evaluated as a predictor of abnormal cranial ultrasound find-

<table>
<thead>
<tr>
<th>TABLE 4</th>
<th>MAP and HOI-33 by Pressor-Inotrope Use and Volume Expander Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>(n)</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>(a) Day 1</td>
<td>Volume expanders ((&gt;10\ mL/kg; N = 81))</td>
</tr>
<tr>
<td>No</td>
<td>33</td>
</tr>
<tr>
<td>Yes</td>
<td>48</td>
</tr>
<tr>
<td>Pressor-inotropes ((N = 82))</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>42</td>
</tr>
<tr>
<td>Yes</td>
<td>40</td>
</tr>
<tr>
<td>(b) Day 2 ((N = 78))</td>
<td>Volume expanders ((&gt;10\ mL/kg))</td>
</tr>
<tr>
<td>No</td>
<td>50</td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
</tr>
<tr>
<td>Pressor-inotropes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>30</td>
</tr>
<tr>
<td>Yes</td>
<td>48</td>
</tr>
<tr>
<td>(c) Day 3</td>
<td>Volume expanders ((&gt;10\ mL/kg; N = 59))</td>
</tr>
<tr>
<td>No</td>
<td>34</td>
</tr>
<tr>
<td>Yes</td>
<td>25</td>
</tr>
<tr>
<td>Pressor-inotropes ((N = 60))</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>23</td>
</tr>
<tr>
<td>Yes</td>
<td>37</td>
</tr>
<tr>
<td>(d) Days 1 through 3 ((N = 84))</td>
<td>Volume expanders ((&gt;10\ mL/kg))</td>
</tr>
<tr>
<td>No</td>
<td>21</td>
</tr>
<tr>
<td>Yes</td>
<td>63</td>
</tr>
<tr>
<td>Pressor-inotropes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>24</td>
</tr>
<tr>
<td>Yes</td>
<td>60</td>
</tr>
<tr>
<td>(e) Days 1 through 3 ((N = 84))</td>
<td>No. of days on volume expanders ((&gt;10\ mL/kg))</td>
</tr>
<tr>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>No. of days on pressor-inotropes</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
</tr>
</tbody>
</table>

For individual days \((a–c)\), mean MAP and HOI-33 are based on BP measurements taken on that day. For overall use \((d)\) and number of days \((e)\), mean MAP and HOI are based on the entire 3-day period. IQR indicates interquartile range (25th–75th percentiles).
ings. Only GA, 5-minute Apgar, minimum pH, minimum Po2, maximum PCO2, minimum hemoglobin, minimum platelet count, and maximum temperature met this criterion. We then repeated the comparisons between abnormal cranial ultrasound findings and each of HOI and mean MAP over days 1 through 3 after stratifying by each of these covariates. In all cases, the stratified P values were similar to the crude (unadjusted) P values, suggesting that none of these variables seemed to exert a confounding effect on the relationship between BP and abnormal cranial ultrasound finding.

To assess the level of illness severity at the onset of these studies, we performed the SNAP-II during the first 12 hours after birth. The mean (± SD) SNAP-II of 23.2 (± 12.4) underscores the high level of illness severity in our population.28 The modified SNAP-II (using the same population mean MAP score for all infants in the BP item of SNAP-II) has been defined. Like others,11,12,27,45 we found a statistically significant difference in MAP over days 1 through 3 between patients with abnormal cranial ultrasound findings. Modified SNAP-II has been shown to be strongly associated with high modified SNAP-II (P = .03) and a significantly lower overall mean MAP over days 1 through 3 (35.9 vs 38.5 mm Hg; P = .007). When SNAP-II and modified SNAP-II were assessed for their relationships with cranial ultrasound outcomes (Table 5), abnormal outcomes were more strongly associated with high modified SNAP-II (P = .01–.05) than with the unmodified SNAP-II (P = .05–.25), suggesting that the association is driven by factors other than BP. Finally, we assessed the relationship between BP and cranial ultrasound outcomes separately in patients with high versus low modified SNAP-II (Table 6). As with GA, for which this association was found only in infants with older GA, the association between BP and abnormal cranial ultrasound findings was present only in infants with lower illness severity.

**DISCUSSION**

Hypotension is frequently diagnosed in the early days after preterm birth.1–3 It is not surprising that a number of studies have associated hypotension with abnormal cranial ultrasound findings.11,12,27,30 Several mechanisms have been proposed to underlie this association. Decreases in BP within the range of functioning cerebral pressure autoregulation trigger cerebral vasodilation, thereby increasing cerebral blood volume. When BP falls below the lower threshold of the autoregulatory plateau, there are 2 major risks for the brain. First, there is a significantly increased risk for cerebral hypoperfusion and ischemia.31–33 Second, the cerebral pressure passivity that results exposes the cerebral vasculature to greater variability in blood flow with changes in BP.34–37 Together, these mechanisms are known to mediate cerebrovascular injury through hypoperfusion-reperfusion mechanisms, as demonstrated in earlier animal studies.38,39 Preventing these developments is a major impetus behind the current practice in most NICUS,5,11,12,23,30 including our own, of diagnosing and promptly treating hypotension in the preterm infant. However, the lower threshold of the autoregulatory plateau or even the existence of such a plateau in the preterm infant remains a contentious issue.40–44

In this study of prolonged high-frequency aortic BP recordings in sick preterm infants during the first 3 days of life, we demonstrated that none of the more commonly used clinical definitions of hypotension reliably predicted early abnormal cranial ultrasound findings. Unlike previous studies that focused on relatively limited aspects of BP, we evaluated the role of BP on early cranial ultrasound outcome in several different ways. Specifically, we explored the effects of MAP variability, mean MAP during 5-minute epochs, and mean MAP during the study period, as well as the cumulative duration of hypotension expressed as the HOI for each definition. Like others,11,12,27,45 we found a statistically significant increase in MAP with increasing GA and postnatal age; however, the 2- to 3-mm Hg increase in MAP during the first 72 postnatal hours was substantially less than that previously reported.12,23,27 This more modest

---

**TABLE 5** SNAP-II for the First 12 Hours of Life and Cranial Ultrasound Outcome

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Early Cranial Ultrasound Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IVH</td>
</tr>
<tr>
<td>SNAP-II</td>
<td></td>
</tr>
<tr>
<td>Low (0–20)</td>
<td>10 of 34 (29%)</td>
</tr>
<tr>
<td>High (≥20)</td>
<td>21 of 49 (43%)</td>
</tr>
<tr>
<td>P</td>
<td>0.25</td>
</tr>
<tr>
<td>Modified SNAP-II</td>
<td></td>
</tr>
<tr>
<td>Low (0–20)</td>
<td>17 of 57 (30%)</td>
</tr>
<tr>
<td>High (≥20)</td>
<td>14 of 26 (54%)</td>
</tr>
<tr>
<td>P</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Modified SNAP-II uses the same population mean MAP score for all infants in the BP item of SNAP-II.

**TABLE 6** AUROC and P Value for Testing Association Among HOI-33, Mean MAP, and Cranial Ultrasound Outcome Stratified by Modified SNAP-II

<table>
<thead>
<tr>
<th>Modified SNAP-II</th>
<th>AUROC Values per Early Cranial Ultrasound Outcome (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IVH</td>
</tr>
<tr>
<td>HOI-33 ≤20</td>
<td>0.73 (0.06)</td>
</tr>
<tr>
<td>HOI-33 ≥20</td>
<td>0.43</td>
</tr>
<tr>
<td>Mean MAP over days 1–3 ≤20</td>
<td>0.71 (0.01)</td>
</tr>
<tr>
<td>Mean MAP over days 1–3 ≥20</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Only P < .01 is shown. AUROC values of >0.50 suggest an association between low BP and abnormal cranial ultrasound findings. Modified SNAP-II uses the same population mean MAP score for all infants in the BP item of SNAP-II.
postnatal increase in MAP was largely attributed to higher starting BPs on the first day, possibly as a result of more liberal early initiation of pressors-inotropes in our unit compared with other centers.23,27

The hypotension thresholds tested in this study either are based on statistical normality in large populations12,46-48 or originate from populations with levels of systemic morbidity different from ours.12,23,27,49; therefore, we tested an MAP threshold at the 10th percentile (ie, 33 mm Hg) for our own cohort. Although this threshold predicted abnormal cranial ultrasound findings in infants between 27 and 30 weeks’ gestation and those with lower illness severity (by the SNAP-II),25 it failed to do so in the smallest, sickest infants (ie, those most vulnerable to brain injury).

We also assessed the effect of rapid BP changes by examining the association between BP variability and abnormal cranial ultrasound findings. Several previous studies have shown an association between BP variability and GM-IVH.1,11,23,27,50-52 Watkins et al found higher BP variability in infants with GM-IVH, but this was not statistically significant. Although the average BP variability did not predict GM-IVH in the study by Bada et al,11 longer durations of high or low variability were predictive of GM-IVH. However, like others,12,13 we found no significant association between BP variability and early abnormal cranial ultrasound findings.

Because appropriate BP targets for intervention remain poorly defined in preterm infants,1,5,53,54 it is not surprising that pressor-inotrope practices for this population vary widely between and even within centers.1,3,54 Despite 3 decades of pressor-inotrope use in the NICU, there is still no direct evidence that current treatment for hypotension decreases mortality or morbidity in the preterm infant.1,5,53 Several authors have challenged the current approaches to management of hypotension in the preterm newborn and, on the basis of an increase in adverse outcome among infants with treated hypotension, have suggested that treatment may contribute to brain injury.1,4,5 In our study, volume expander use after the first day was associated with abnormalities on cranial ultrasound, but no significant association was seen with pressor-inotrope use. As in previous studies,4,43 infants who received volume expanders or pressors-inotropes had significantly lower BPs. In fact, the HOI-33 measures in Table 4 show that infants who were treated with pressors-inotropes had mean MAP values during the study period that were below our population’s 10th percentile >15-fold longer than untreated infants. Although these observations may of course reflect confounding by indication, the prompt treatment of hypotension in our NICU, the expected rapid response to pressors-inotropes, and the duration of treatment (58% of treated infants received pressors-inotropes for >1 day) suggest other possible explanations. First, it is possible that infants who are prone to hypotension are poorly responsive to current therapies. Second, the BP targets pursued or the vigor of that pursuit may be inadequate.

There are several potential explanations for why we, unlike others,11,12,27,30 found no significant association between currently applied definitions of hypotension and early abnormal cranial ultrasound findings. First, the severity of illness in our population, as indicated by the high SNAP-II, may be higher than that in previous studies;23,27; therefore, factors other than hypotension may have contributed to the development of abnormal cranial ultrasound findings. Consistent with this, the modified SNAP-II, which did not account for BP, was more closely associated with abnormal cranial ultrasound findings than the unmodified score. Second, it is possible that our techniques for BP measurement and analysis were not sufficiently sensitive to detect an association. Previous studies were limited by small patient numbers30,46,55 or differed from our study in their methods of BP measurement and analysis.12,23,27,45 The techniques for BP measurement in previous studies12,23,30 have ranged from isolated measurements taken hourly to every 24 hours, to BP averaged over extended periods. Our goal was to capture BP behavior across a broad temporal spectrum by using high-frequency measurements and then generating measures of rapid, short-term, and long-term BP changes; however, despite this comprehensive analysis of MAP, our recordings were not continuous throughout the 3 days, and it is possible that injurious hemodynamics that are responsible for abnormal cranial ultrasound findings may have gone undetected during off-hours in our study.

Another potential reason that early life BP changes were not predictive of abnormal cranial ultrasound findings in our study may relate to the difficulty of making a precise temporal association between BP changes and injury. It is possible that hemodynamic changes that are responsible for brain injury in the preterm infant are brief and intermittent. Because the onset of brain injury in sick preterm infants is usually clinically subtle or silent and clinical protocols usually separate cranial ultrasound by at least days, injurious short-term hemodynamic changes may be averaged out during prolonged recordings during the high-risk period for brain injury. We used cranial ultrasound as our principal outcome measure12,23,27 because it remains the only bedside imaging technique currently available and therefore the only technique for brain imaging during the acute period of critical illness in sick preterm infants. Cranial ultrasound reliably detects the onset of hemorrhage but is both insensitive and delayed in its detection of hypoxic-ischemic injury;56-58 therefore, although it is likely that we captured major focal parenchymal lesions as echodensities on cranial ultrasound, it is possible that lesser and more diffuse degrees of white matter injury went undetected. Furthermore, because the cranial ultrasound in
our study was clinically indicated, it was not performed at consistent times; however, all infants had ultrasound studies between 5 and 10 days, and because this period is beyond the greatest risk for especially GM-IVH, we based our cranial ultrasound outcome on this study.

Recent studies have suggested that BP may not reliably reflect brain perfusion during critical early periods in the transition from fetal to premature neonatal hemodynamics,59–62 These studies used superior vena cava blood flow as a surrogate for cerebral blood flow, and because the validity of this technique is untested, the relevance of a pressure-flow disconnect to our findings remains unclear. Finally, other mechanisms of brain injury, such as those mediated by circulating inflammatory substances63–68 or abnormalities in circulating carbon dioxide,69–73 have been proposed as principal mediators of brain injury in the preterm infant. These issues remain unresolved and in need of additional study.

One major factor that is overlooked in the debate about the relation between BP and brain injury in preterm infants is the dynamic yet tenuous nature of intrinsic cerebrovascular control in these patients.17 The interaction between immature systemic and cerebral hemodynamics during the period of circulatory transition is likely to be extremely complex and remains very difficult to study continuously at the bedside of sick preterm infants. In a recent study,17 we used continuous direct measurements of BP time-locked to continuous near-infrared spectroscopy measures of cerebral hemodynamics to demonstrate a high prevalence and fluctuating nature of cerebral blood flow pressure passivity in sick preterm infants. Most important, the relationship between cerebral pressure passivity and BP was unpredictable, with pressure passivity occurring during both normal and hypotensive BP levels (as currently defined) in the same infant over time.17 These findings emphasize the dynamic nature of the cerebral pressure autoregulatory plateau and suggest that factors other than BP may influence the BP thresholds for cerebral pressure passivity. These questions are under investigation.

CONCLUSIONS
In this study we failed to find an association between various widely used definitions of hypotension in preterm infants and the emergence of early brain injury as detected by cranial ultrasound. This lack of association was particularly true of the smallest, sickest, and therefore most vulnerable infants. These findings suggest that adherence to current population-based guidelines for the diagnosis and management of hypotension might not be effective in preventing brain injury in sick preterm infants. On the basis of our previous work,17 we propose that this lack of association is attributable to the influence of factors other than BP on cerebral autoregulation. Without the ability to measure cerebrovascular responses continuously, identifying and applying optimal BP levels to prevent brain injury across whole populations of sick infants or even in the individual infant may be an exercise in futility.

ACKNOWLEDGMENTS
This work was supported by National Institutes of Health grant PO1NS38475, the LifeBridge Fund, the Caroline Levine Foundation, the Trust Family Foundation, and a postdoctoral fellowship (to Dr Limperopoulos) from the Canadian Institutes of Health Research.

We thank Shaye Moore for assistance with manuscript preparation and the patients and families for participation in this study.

REFERENCES
16. Kluckow M, Evans N. Low superior vena cava flow and intra-
34. Lassen NA. Cerebral blood flow and oxygen consumption in man. Physiol Rev. 1959;39:183–233
57. Hope PL, Gould SJ, Howard S, Hamilton PA, Costello AM, Reynolds EO. Precision of ultrasound diagnosis of pathologi-
cally verified lesions in the brains of very preterm newborn infants. Dev Med Child Neurol. 1988;30:457–471

**RANDOMIZED CLINICAL TRIALS**

*Pediatrics* requires investigators to register their clinical trials in a public trials registry. The members of the International Committee of Medical Journal Editors (ICMJE) will consider most clinical trials for publication only if they have been registered (see *N Engl J Med*. 2004;351:1250–1251). Current information on requirements and appropriate registries is available at www.icmje.org/faq.pdf.
Impact of Singular Excessive Computer Game and Television Exposure on Sleep Patterns and Memory Performance of School-aged Children

Markus Dworak, DiplSportwiss (MSc), Thomas Schierl, PhD, Thomas Bruns, PhD, Heiko Klaus Struder, PhD

Institute of Motor Control and Movement Technique and Institute of Sports, Media and Communications, German Sport University Cologne, Cologne, Germany

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Television and computer game consumption are a powerful influence in the lives of most children. Previous evidence has supported the notion that media exposure could impair a variety of behavioral characteristics. Excessive television viewing and computer game playing have been associated with many psychiatric symptoms, especially emotional and behavioral symptoms, somatic complaints, attention problems such as hyperactivity, and family interaction problems. Nevertheless, there is insufficient knowledge about the relationship between singular excessive media consumption on sleep patterns and linked implications on children. The aim of this study was to investigate the effects of singular excessive television and computer game consumption on sleep patterns and memory performance of children.

METHODS. Eleven school-aged children were recruited for this polysomnographic study. Children were exposed to voluntary excessive television and computer game consumption. In the subsequent night, polysomnographic measurements were conducted to measure sleep-architecture and sleep-continuity parameters. In addition, a visual and verbal memory test was conducted before media stimulation and after the subsequent sleeping period to determine visuospatial and verbal memory performance.

RESULTS. Only computer game playing resulted in significant reduced amounts of slow-wave sleep as well as significant declines in verbal memory performance. Prolonged sleep-onset latency and more stage 2 sleep were also detected after previous computer game consumption. No effects on rapid eye movement sleep were observed. Television viewing reduced sleep efficiency significantly but did not affect sleep patterns.

CONCLUSIONS. The results suggest that television and computer game exposure affect children’s sleep and deteriorate verbal cognitive performance, which supports the hypothesis of the negative influence of media consumption on children’s sleep, learning, and memory.
In Western industrial countries, television and computer game consumption take up a large part of children’s time awake. The proportion of television users among children who are aged 9 to 16 years ranges from 98% to 100%. Some studies have indicated that excessive television and video game consumption could result in psychiatric symptoms such as aggressive behavior, attention problems, hyperactivity, scholastic problems, and somatic complaints. Despite this knowledge, average media consumption of the children in this age group is >4 hours/day. Furthermore, these sedentary activities are frequently associated with significant behavioral consequences, including decreased physical activity and physical fitness, poor eating habits, and obesity, and could impair development in childhood and adolescence.

Despite the enormous progress in media research, there is insufficient knowledge about the effects of singular excessive media exposure on behavioral states and sleep patterns in children. Most examinations focused on long-term effects of media exposure on children’s well-being, and only a few studies examined the effects of television and computer consumption on children’s sleep quality. Negative effects on sleeping behavior, such as sleep-onset delay, night waking, sleep anxiety, and shortened sleep duration, were observed in recent studies and suggest that media exposure could impair sleep quality. Finally, watching television ≥3 hours/day during adolescence elevates the risk for frequent sleep problems in early adulthood.

Sleep is essential for children’s health and development and possibly plays an important role in learning and memory. Neuroscientific theories support the notion that emotions could influence learning processes. Especially recently acquired knowledge is very sensitive in the subsequent consolidation period, and what the children emotionally experience within the hours after learning influences it decisively. The influence of interactive media consumption on emotions was also established. Thus, it could be hypothesized that television and computer game consumption after a learning period could impair memory consolidation and performance. The purpose of this polysomnographic study was to evaluate direct short-term effects of singular excessive television and computer game consumption on sleeping patterns and cognitive performance of school-aged children.

METHODS

Participants
Eleven male children volunteered to participate in this study (age: 13.45 ± 1.04 years; height: 1.64 ± 0.09 m; weight: 48.23 ± 5.97 kg; BMI: 17.82 ± 1.29 [all means ± SD]). Children and their parents for this study were volunteers from 8599 families who previously were recruited for and participated in the Healthy Sleep for Cologne Children study, an epidemiologic study of sleeping behavior and sleep complaints in children. All participants were junior high school children in the first grade. Participants were selected in random order after fulfillment of the following criteria: age between 12 and 14 years, great health, no medications, German nationality, and male gender. Only male participants were selected for this study because previous studies indicated that menstrual cycle phase as well as oral contraceptive could influence sleeping behavior. In precocious girls, these factors could influence nocturnal sleep. In addition, media research studies indicate that boys spend more than twice as much time playing video games as girls do; therefore, video game playing is a greater impact factor in boys’ leisure time. After consideration of the mentioned criteria, 3580 children were contacted with detailed information regarding the study setting; a total of 1321 consenting children gave a response rate of 36.9%. All participants and parental authorities signed informed consent forms and completed a medical questionnaire before the experiment. They were informed that they could quit the study at any stage. All participants showed good health status and had no sleep complaints. Participants were instructed to refrain from additional physical activity, passive body warming (eg, taking a warm bath or shower), and sleep during the day on the experimental days. They did not consume any kinds of caffeine, nicotine, or alcohol, and they were not exposed to a large stress load.

Experimental Design
Each participant underwent 3 investigation days in a randomized, crossover manner. The interval between the experiments was exactly 1 week. On 2 different experimental days, children were exposed to 2 types of media. On 1 occasion, they played interactively a computer game (Need for Speed—Most Wanted; Electronic Arts, Redwood City, CA) for 60 minutes. Recent studies suggest that this is nearly the average time per day that children spend playing computer games. Furthermore, the participants watched a subjectively exciting video film on television. Every participant could choose from among 3 films (Harry Potter and the Prisoner of Azkaban, Star Trek: Nemesis, and Mary Higgins Clark’s Love Music, Loves to Dance) but was not allowed to have seen the film before the experimental day. Media exposure occurred between 6:00 pm and 7:00 pm (2–3 hours before bedtime). This time of day was selected because the children have usually finished their homework at that time and start their leisure time. Under control conditions, the participants adhered to their normal daily patterns but were not allowed to watch television or play computer games.

Four to 5 hours before bedtime, when children usually did their homework, a visual and verbal memory...
test (VVM) was conducted on each experimental day (Swets Test Services, Frankfurt am Main, Germany). The test served to determine the short-term and longer-term memory of visuospatial and verbal materials and was subdivided into 2 subtests. For examination of the memory performance in the visually spatial area, a map with a marked path was shown to the children for 2 minutes. Immediately after this, the participants were asked to draw the path on such a map from memory. The second subtest served to examine the memory of facts. In this case, text that included names, numbers, and terms was presented for 2 minutes. After the presentation, the facts were immediately (T1) asked for in writing. Both tests were repeated after each experimental day within a 24-hour interval (T2) without renewed attraction. The raw analyzed results of the 2 subtests were intended for the 2 test times. A calculation of the loss of memory performance for every subtest was conducted by the following formula: \( T1 - T2 = \frac{(T2 - T1)}{T1} \times 100. \)

**Data Recording and Analysis**

Before participant went to bed, polysomnographic measurements were conducted using a portable sleep data recorder (Variopost-SLP 2.0; Becker Meditec, Karlsruhe, Germany). An expert affixed the electrodes between 7:30 PM and 8:30 PM and removed them after participants awoke in the morning. The participants were also instructed to adhere to their normal evening routines and to get up at their usual time. General bedtime was between 8:30 PM and 9:30 PM, when the room light was turned off. Morning waking was between 6:00 AM and 7:00 AM. During the study nights, the participants slept in their own homes and always under the same timing and temperature conditions to standardize sleeping conditions for each participant. They were also instructed to adhere to their normal evening routines and fill in a sleep diary. The sleep diary contains questions about daytime sleepiness, subjectively rated sleepiness, and subjectively rated awakenings. An adaptation night was assigned on the day before the experiment to reduce the possible “first-night effect.”

The monitoring montage consisted of 3 electroencephalograph channels (C3-A2, C4-A1, and Oz-A2), bilateral electrooculograph, and submental chin electromyography. The measuring procedure followed the standards for performance and evaluation of polysomnographic studies of the pediatric group in the German Sleep Society.28

Each polysomnograph was scored by Somnolyzer 24x7 (Siesta Group, Vienna, Austria). The system included a raw data quality check, a feature extraction algorithm (density and intensity of sleep/wake-related patterns, eg, sleep spindles, \( \delta \) waves, slow eye movements, and rapid eye movements), a feature matrix plausibility check, a classifier designed as an expert system, and a rule-based smoothing procedure for start and ending of stages. In addition, a structured quality control by 2 experts including a visual correction was accomplished. Studies showed that 2 Somnolyzer 24x7 analyses revealed an inter-rater reliability close to 1, representing an overall agreement of 99.4% (Cohen’s \( \kappa \): 0.991). This confirms that the variability induced by the quality control procedure, whereby \( \sim 1\% \) of the epochs are changed, could be neglected.29

For each polysomnograph, a number of measurements of sleep architecture and sleep continuity were derived. Measurements of sleep architecture included minutes and percentage of total sleep time (%TST) of stage 1 sleep, minutes and %TST of stage 2 sleep, minutes and %TST of stage 3 sleep, minutes and %TST of stage 4 sleep, minutes and %TST of slow-wave sleep (SWS; stages 3 + 4 sleep), and minutes and %TST of rapid eye movement (REM) sleep. Continuity measurements consisted of TST, sleep-onset latency (SOL), latency of stage 1, latency of stage 2, latency of stage 3, latency of stage 4 and REM sleep, wake time after sleep onset and sleep efficiency.

**Statistical Analysis**

Statistical analyses for significant differences of the natural sleep cycle data were performed by using repeated-measurements analysis of variance and Bonferroni test as a posthoc test. A paired \( t \) test was used to analyze vigilance test (VVM) data. We used SPSS 12.0 (SigmaStat Statistical Software, Chicago, IL) and Statistica 7.1 software (StatSoft, Tulsa, OK) for Microsoft Windows. The significance level of all statistical tests was set at \( P < .05 \).

**RESULTS**

All sleep parameters showed normal distributions. Because 1 of the children was dyslexic, only the tests and polysomnographs of the 10 healthy participants were analyzed. TST remained unchanged among the 3 experimental nights. Mean TST was 511.80 \( \pm \) 44.44 minutes. The results showed a significant \( (P < .05) \) decrease in sleep efficiency after only television exposure (Table 1). SOL increased significantly \( (P < .05) \) after computer game stimulation compared with basal conditions. Under baseline conditions, SOL was 10.83 \( \pm \) 8.33 minutes. Significant increases to 32.50 \( \pm \) 25.67 minutes \( (P < .05) \) were detected after computer game playing. No effect was found after television exposure. Also, significant prolongations in latent periods to stage 2 and stage 4 were observed only after interactive computer game playing \( (P < .05) \).

Table 2 provides the means and SDs of the participants for sleep-architecture parameters on experimental days. Especially computer game playing resulted in a shift of sleep stages (Fig 1). Participants spent significantly \( (P < .05) \) more time in sleep stage 2 compared with basal conditions (Fig 2). Case-wise data showed
considerable increases over 50.0 minutes of stage 2 sleep in 7 children. In contrast, after television exposure, none of the participants showed increases 50.0 minutes in stage 2 sleep compared with basal conditions. Furthermore, percentage distribution of sleep stages showed a significant decrease in SWS after computer game consumption related to basal conditions ($P < .05$; Fig 3).

Also, a more detailed analysis of case-wise data supports the explicit effects of computer game consumption on SWS. Declines of SWS $>5.0\%$ were observed in 7 participants; 2 of these reached $>10.0\%$ (10.02% and 13.27%). Conversely, only 1 participant showed decreased SWS proportions $>5.0\%$ after television consumption compared with basal conditions. There were no significant changes in either the absolute duration or the relative proportion of REM sleep. Also, the evaluation of the sleep diary resulted in no significant differences among the experimental days (Fig 3).

**Visual and Verbal Memory Test**

The evaluation of the visual and verbal memory test yielded a negative influence of the computer game exposure on the verbal memory for facts. Computer game playing led to a significant ($P < .01$) decline of the verbal memory performance ($-46.83 \pm 18.32$) compared with basal conditions ($-18.09 \pm 24.78$). In addition, case-
wise data showed declines >20.0% in 8 participants. In contrast, no significant changes were observed after television exposure after which only 2 participants showed a decline of verbal memory performance >20.0%. Cognitive performance in the visuospatial area did not differ among the experimental days (Fig 4).

DISCUSSION
This study demonstrates that singular excessive media exposure affects children’s sleep architecture, sleep continuity, and verbal memory performance. Particularly, interactive computer game consumption resulted in prolonged SOL, more sleep time in stage 2, less SWS as a percentage of TST in subsequent sleep, and declines in verbal memory performance; therefore, our results provide supplementary evidence for a negative influence of excessive media consumption on children’s sleep, health, and performance.

Most previous studies observed long-term term correlations between television viewing, computer game playing, and Internet use and general health problems, whereas our study proves singular excessive short-term effects; nevertheless, our findings are consistent with previous studies and support that there is a negative influence of excessive media consumption on health and well-being. Excessive media consumption was associated with an elevated risk for psychiatric and social problems such as aggressive behavior, attention problems, hyperactivity, and scholastic problems. Also, links have been suggested with certain somatic complaints resulting from the sedentary execution of media consumption in children’s leisure time, including decreased levels of physical activity and physical fitness, poor eating habits, and an increased risk for obesity.

Unfortunately, only a few studies have assessed the effects of media consumption on children’s sleep. Recent results showed increased SOL, an elevated risk for midnight waking, difficulties in falling asleep, and a reduced sleep quality after television and computer game consumption. Poor sleep quality was associated with mental health problems, inferior school performance, and somatic complaints. Sleep difficulties were significantly associated with both behavioral problems, such as school attendance problems, and higher levels of tiredness. It has also been found that excessive television viewing may be connected with diverse sleep disturbances during adolescence. The presence of a television in the child’s bedroom results in significant modifications of sleep–wake parameters, especially related to bedtime and sleep duration, and thus is the most powerful predictor of overall sleep disturbance and bedtime resistance.

To our knowledge, only 1 study previously examined the effects of media consumption (computer game playing) on sleep patterns in children. In accordance with Higuchi et al, we used polysomnographic measurements to define sleep stages and sleep latencies. In both studies, a significant increased SOL after singular excessive computer game playing was observed compared with control conditions. Contrary results were detected in sleep stages. Whereas Higuchi et al noticed less REM sleep and no changes in SWS after computer games, we detected more sleep in stage 2 and reduced amounts of SWS but no effects on REM sleep. The decrease in SWS after computer game playing in this
study may reflect children’s high arousal state. Differences in the age of the participants, type of computer game, place and type of polysomnographic measurements, and sleep time may be reasons for the different observation.

Possibly, different claims of television and computer game consumption on the central nervous system were decisive for alterations of children sleep. Previous studies showed different effects of television viewing and video game playing on several physiologic parameters. Unlike television viewing, which expends the same energy as sitting quietly, interactive video game consumption results in significant increases in various physiologic and metabolic variables in young children, including heart rate, blood pressure, respiratory rate, energy expenditure, and ventilation, and thus a higher arousal state of the central nervous system. The magnitude of these changes was below standard physical exercise and national health recommendations and did not affect metabolic, cardiovascular, pulmonary, homodynamic, and endocrine systems in the whole body and the brain as physical activity does. A higher arousal state within the hours before sleeping could influence subsequent sleep.

Another interesting finding was a significant decline of verbal memory performance after computer game playing compared with basal conditions. Modern neuroscientific theories support the notion that strong emotional experiences, such as computer games and thrilling films, could decisively influence learning processes. Because recently acquired knowledge is very sensitive in the subsequent consolidation period, emotional experiences within the hours after learning could influence memory consolidation considerably. Interactive video games are challenging, sometimes frustrating, exciting, and often surprising, and during playing, individuals may experience a range of emotions accompanied by physiologic changes. In addition, studies with positron emission tomography scans showed a significant release of the neurotransmitters dopamine and norepinephrine in the brain during video game playing. Dopamine as well as norepinephrine are thought to be involved in learning, reinforcement of behavior, emotion, and sensorimotor coordination and thus able to influence memory processing decisively.

Evidence in molecular genetics, neurophysiology, and the cognitive neuroscience supports an important role for sleep in learning and reprocessing of memories. Only a single night of restricted sleep led to impaired cognitive functions, such as abstract thinking and verbal creativity in children. Presumably, both REM and SWS are involved in the consolidation process, in which SWS is particularly favorable to explicit memory traces. It has been proposed that during SWS, the lower acetylcholine levels facilitate the transmission of information from the hippocampus back to the cortex. High acetylcholine levels during REM sleep would allow the neocortex to undergo a process of reanalysis and thereby develop new feed-forward representations for behavior. Because computer game exposure resulted in likewise reduced amounts of SWS, our observations support the hypothesis for the role of SWS in explicit memory consolidation.

Video viewing did not affect visuospatial and verbal memory performance. It could be assumed that the media content is responsible for subsequent effects on sleep and memory performance. Afterward, none of the participants judged the chosen film as very thrilling, indicating that the thrilling factor of the selected films was not high enough for the children. Results showing that exposure to adult media content may have a stronger impact than media exposure time supported this notion. In general, just 13% of the young people in this age group have parental control with rules about the content of their media consumption. Especially adult (violent/sexual) media content and associated individual excitement could affect sleep and learning in children.
It could be hypothesized that media exposure influences memory processing in decisive ways: temporary through emotional influences on the consolidation process as well as disrupted SWS and in the long-term by a chronic diminution of physical activity. An inverse relationship between time spent using video games and daily physical activity has already been observed. Positive effects of physical exercise on brain structures, functions, and memory processing were examined in recent studies and supported by cross-sectional observations that showed a positive association between physical activity and academic and accompanied improvements of concentration and classroom behavior. Finally, our results could provide a plausible explanation concerning the effects of media exposure on poor school completion and especially poor reading skills, derived from the results of the VVM.

This study supports the influence of singular excessive media consumption on both sleep and cognitive performance in children. In addition, the impact of media on children’s health and well-being is widely recognized and considered a serious problem in modern society. Our results were supported by previous studies that observed that movie, television, and video game use during the middle school years was uniformly associated with a negative impact on school performance. Also, children with lowest grades spend more time playing video games and less time reading than those with the best grades. Additional examinations confirmed that excessive television viewing in early childhood was associated with a higher subsequent risk for development of attention-deficit/hyperactivity disorder (ADHD). Limiting young children’s exposure to television as a medium during formative years of brain development may reduce children’s subsequent risk for developing ADHD and social and scholastic problems. Children spend one third of their life in sleep, indicating the importance of sleep for children’s development and health; therefore, the negative effects of excessive media consumption may be a significant concern. Our findings add experimental support to the importance of parental limits on media content and time. Additional work is needed to determine relationships among time, content, and type of media affecting children’s health, sleep, and memory and to give advice for useful contact with entertainment media.

CONCLUSIONS
Our data indicate that excessive media consumption, especially computer game playing, impairs sleep patterns and verbal cognitive performance in children. Because children’s sleep-related problems seem to be highly persistent; prevalent; and associated with somatic complaints, psychiatric symptoms, especially behavioral and emotional symptoms, attention problems such as hyperactivity, and scholastic problems, they constitute a considerable and growing health problem among children and therefore should receive more attention. This study demonstrates that more effort should be directed to screening sleep disturbances after media consumption, helping parents to perceive the negative effects of media consumption on health and sleep and to provide adequate guidance for their children when needed.

ACKNOWLEDGMENTS
This work was supported by the Institute of Motor Control and Movement Technique and the Institute of Sports, Media and Communications and performed at the Institute of Motor Control and Movement Technique (German Sport University Cologne).

REFERENCES


Associations Between Content Types of Early Media Exposure and Subsequent Attentional Problems

Frederick J. Zimmerman, PhD, Dimitri A. Christakis, MD, MPH
Child Health Institute and Departments of Health Services and Pediatrics, University of Washington, Seattle, Washington; Seattle Children’s Research Institute, Children’s Hospital and Regional Medical Center, Seattle, Washington

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Television and video/DVD viewing among very young children has become both pervasive and heavy. Previous studies have reported an association between early media exposure and problems with attention regulation but did not have data on the content type that children watched. We tested the hypothesis that early television viewing of 3 content types is associated with subsequent attentional problems. The 3 different content types are educational, nonviolent entertainment, and violent entertainment.

METHODS. Participants were children in a nationally representative sample collected in 1997 and reassessed in 2002. The analysis was a logistic regression of a high score on a validated parent-reported measure of attentional problems, regressed on early television exposure by content and several important sociodemographic control variables.

RESULTS. Viewing of educational television before age 3 was not associated with attentional problems 5 years later. However, viewing of either violent or nonviolent entertainment television before age 3 was significantly associated with subsequent attentional problems, and the magnitude of the association was large. Viewing of any content type at ages 4 to 5 was not associated with subsequent problems.

CONCLUSIONS. The association between early television viewing and subsequent attentional problems is specific to noneducational viewing and to viewing before age 3.
Poor attention regulation among grade-school children significantly impairs educational performance, imposes significant cost burdens on schools, and is a source of considerable anxiety for parents and teachers. When manifested as attention-deficit/hyperactivity disorder (ADHD), such problems are among the most common chronic diseases of childhood, affecting somewhere between 4% and 11% of children. Although the genetic role in the cause of ADHD has been well established, comparatively little is known about the environmental risk factors, yet given an increasing recognition of gene-environment interactions in the genesis of ADHD, more research into environmental factors is clearly warranted. As stated in the Surgeon General’s report on mental health, “For most children with ADHD, the overall effects of these gene abnormalities seem small, suggesting that nongenetic factors also are important.” There has been speculation of a role for television in ADHD for decades, and although much of this speculation has been unscientifically based, there is sound reason to suspect a possible role for early television, especially if it is particularly heavy and occurs early in life. The first 3 years of life are characterized by tremendous development, much of which occurs in response to the environment. The newborn brain is developing rapidly and characterized by great plasticity in response to the child’s environment.

The theoretical mechanisms through which early television viewing might impair healthy development of attention regulation may be moderated through the type of on-screen content. The theory of displacement suggests that television’s deleterious effects operate by displacing developmentally appropriate learning opportunities with an attention-grabbing stimulus with little developmental value. In this theory, because educational shows such as Sesame Street and Dora the Explorer are designed to foster learning, they will be less harmful and may even be helpful compared with shows that are produced purely for entertainment. The theory of formal features suggests that the fast pacing and rapid scene changes that are characteristic of television reward fixed attention to a constantly changing stimulus and do not reward self-directed attention to opportunities for learning. Here again, educational shows would be expected to be less damaging because their pacing is typically much slower.

An earlier study found modest but significant associations between total television viewing before age 3 and problems with attention regulation at age 7, controlling for a variety of possible confounders, including the level of parental emotional support and cognitive stimulation provided to the child in the first years of life. Another study found no association between television viewing at age 5 and attention problems at age 6. A third study of 170 children aged 2 to 5 found a positive and significant contemporaneous association between television viewing and teacher report of ADHD symptoms based on the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, and also between television viewing and actigraph measurements of child movement. No study to date has disaggregated viewing by content type. This work raises 2 important questions: (1) Is the association between early television viewing and subsequent attentional problems moderated by the type of content viewed? (2) Are the first 3 years of life a critical age range within which children are more vulnerable to the effects of media than is the case at later ages?

METHODS

Data Source
The Panel Survey of Income Dynamics (PSID) is a longitudinal study begun with 4800 families in 1968 with a variety of funding sources and overseen by the National Science Foundation. In 1997, the Child Development Supplement (CDS), which is a questionnaire administered to the primary caregivers of 3563 children aged 0 to 12, was added. The questionnaire, funded by the National Institute of Child Health and Human Development, included detailed demographic data, psychological and behavioral assessment of parents and children, and time use diary data from 1 randomly chosen weekday and 1 randomly chosen weekend day during a school year (September through May). These diaries include primary and secondary activities during a 24-hour period. Such time diaries have been used extensively in research and have excellent validity when compared with direct observation of activities. Among eligible households in the core PSID sample, 1997 CDS data were obtained for 88%. In 2002, the respondents of the first CDS were followed up with a second, similar CDS (CDS-II). The follow-up rate from the 1997 CDS to the CDS was 91%.

These data are in the public domain. The research was approved by the Children’s Hospital and Regional Medical Center Institutional Review Board.

Children
We included all children who participated in the study, were younger than 5 years, and had follow-up data 5 years later. We divided the sample into children who were younger than 3 and those who were 4 to 5 years of age at baseline (so that at follow-up they were either 5–8 or 9–10 years of age).

Outcome Measure
The PSID included the Behavior Problems Index (BPI), a brief, parent-response inventory of behavior problems that is derived from and similar to the Achenbach Child Behavior Checklist. Our dependent variable used the “hyperactive” scale of the BPI, which has been validated...
and extensively used in published research. It consists of 5 items that each measure elements of attention-regulation: difficulty concentrating, impulsive, easily confused, has obsessions, and restless. Each item allows 3 responses: not true, sometimes true, and often true, coded 1, 2, and 3, respectively. A summary score ranging from 5 to 15 is constructed by summing the scores. This scale has a Cronbach’s α value of .68 and a 2-week test-retest reliability of .68. We dichotomized this scale at ≥11, or 2 SDs above the mean. We chose this cutoff in part because it yielded a prevalence for attentional problems that is similar to published reports of ADHD prevalence among similar-aged children in community samples. In addition to capturing behavioral manifestations (restlessness and impulsivity), it captures the cognitive deficits (obsessions, easily confused, and difficulty concentrating) that characterize problems of attention regulation.27

Main Predictors
We used content data for programming that were developed from the time diaries. A coding system was developed previously to classify television shows and movies on video (hereafter both referred to as “shows”) by content in several dimensions. An educational attribute was assigned when the show had a clear intent to educate, with an explicit cognitive or prosocial component, as follows: Cognitive informative: teaches a lesson with content similar to that found in schools (math skills, reading skills, other school readiness skills). Social informative: teaches a lesson about appropriate behavior or interpersonal interactions (eg, sharing, friendships, drug education).

Violent content was ascribed when “violence was a central and integral part of the plot or of the main characters’ occupations, if the lead characters’ main purpose was to fight or flee from violence, or if there was more violence in the program than would be expected in the everyday life of a child.” Although the term “violent” is used here, it should be understood that the definition includes hostile language, threatening behavior, and cartoon violence as well as realistic violence. A minimum of 2 coders evaluated each show for violence and educational content, with an interrater κ score of 0.81. Differences were resolved by discussion. For this study, we classified shows into 3 categories: educational, nonviolent entertainment (ie, not violent and not educational), and violent entertainment (ie, violent and not educational). No educational shows contained violence.

Some shows could not be coded for violence, either because the name was inadequately reported (eg, “cartoons,” “channel 13”) or because the researchers could not evaluate the violent content of an uncommon but named video or show. These shows, 20% of the total, were included in the “nonviolent” category. Because these shows were not separately flagged, it was not possible to conduct sensitivity analyses by placing these shows into different categories or into their own category. Examples of common shows in each category are presented in Table 1. These example shows were commonly viewed by children aged 0 to 3 and those aged 4 to 5, with little difference in popularity across ages.

Potential Confounders
We adjusted for several child and parental attributes that may confound the relationship between early television viewing and attentional capacity by virtue of being plausibly associated with both. These include the child’s age, race/ethnicity, and gender. Each of these is associated with television viewing and may affect either the risk or expression of symptoms of attentional problems. We also controlled for region and urbanicity of residence, because of the possibility of cultural differences in reporting of symptoms across these attributes. We also controlled for several indicators of socioeconomic adversity that may be associated with early television viewing and that are risk factors for attentional problems, including number of children, mother’s and father’s educational levels, the mother’s score on a depression instrument, presence of father in the household, and parental conflict. The father’s education was available for all children, and the level of parental conflict was imputed with the sample mean when the father was not present. In addition, we controlled for attributes of the child and the family that might affect the amount and kinds of early-life stimuli that the child receives, which in turn could affect development of executive function, which is thought to underlie attentional capacity. These include birth order and validated measures of emotional support and cognitive stimulation. By controlling for these characteristics, we at least partially control for the possibility that children who grow up in a rich cognitive environment are less likely to be exposed to high amounts of early television.

<table>
<thead>
<tr>
<th>Popular Titles According to Content Type</th>
<th>Educational</th>
<th>Nonviolent Entertainment</th>
<th>Violent Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barney</td>
<td>Flintstones (cartoon)</td>
<td>Space Jam</td>
<td></td>
</tr>
<tr>
<td>Sesame Street</td>
<td>Aristocats</td>
<td>Lion King</td>
<td></td>
</tr>
<tr>
<td>Winnie the Pooh</td>
<td>Rugrats</td>
<td>Power Rangers</td>
<td></td>
</tr>
<tr>
<td>Arthur (cartoon)</td>
<td>Babe</td>
<td>Scooby Doo</td>
<td></td>
</tr>
<tr>
<td>Blue’s Clues</td>
<td>Bambi</td>
<td>Looney Tunes</td>
<td></td>
</tr>
<tr>
<td>Doug</td>
<td>Family Matters</td>
<td>America’s Funniest Home Videos</td>
<td></td>
</tr>
</tbody>
</table>
Nonviolent entertainment television viewing was also associated with approximately double the odds for attentional problems 5 years later (odds ratio [OR]: 2.20; 95% confidence interval [CI]: 1.19–4.08). Logistic regression was used to test for the association of early television viewing by content with BPI score indicating attentional problems 5 years later, controlling for demographic attributes and parenting variables as discussed. Appropriate sampling weights were used to provide nationally representative samples. All analyses were conducted by using Stata 9.0 statistical software.19

RESULTS
For children who were aged 0 to 35 months in 1997, 560 children’s parents completed both the CDS-I and the CDS-II. For the children who were aged 4 to 5 years in 1997, the figure is 407. Parents of boys, parents of black children, single mothers, and mothers with lower levels of education were significantly less likely to complete the CDS-II than other parents. No category of television viewing was associated with failure to complete the CDS-II. There were no other statistically significant differences in any of the 1997 covariates between those whose parents did versus did not complete the CDS-II. Of the 560 younger children, complete data were available for 542. Of the 407 older children, 391 had complete data. The children with missing data for any covariate were dropped.

Table 2 presents descriptive statistics of television viewing and the outcomes. Table 3 presents the results of fully adjusted logistic regressions of attentional problems on television viewing by content type for the 2 age groups.

For children who were younger than 3 years, educational television was not significantly associated with subsequent attentional problems. By contrast, each hour per day of average viewing of violent entertainment television was associated with approximately double the odds for attentional problems 5 years later (odds ratio [OR]: 2.20; 95% confidence interval [CI]: 1.19–4.08). Nonviolent entertainment television viewing was also significantly associated with subsequent attentional problems, although not as strongly (OR: 1.73; 95% CI: 1.02–2.94). Figure 1 presents these results graphically. For children aged 4 to 5 years, television viewing of any content type was not significantly associated with attentional problems 5 years later.

DISCUSSION
This analysis provides a replication and enhancement of previous studies of the association between early television viewing and subsequent development of symptoms of attentional problems. The results here are consistent with a previous study that found a significant association between television viewing before age 3 and subsequent attentional problems15 and also with a previous study that failed to find a significant association between television viewing at age 5 and subsequent attentional problems.16 The replication of these 2 distinct results in 1 data set suggests that it is viewing specifically before age 3 that is relevant.

This analysis also enhances the insights from these previous studies by isolating the independent effects of 3 different types of media content: educational shows (eg, Sesame Street), nonviolent entertainment shows (eg, The Aristocats), and violent entertainment (eg, Looney Tunes). Viewing violent or nonviolent entertainment before age 3 is shown to be a significant and meaningful risk for the
development of subsequent attentional problems, whereas viewing educational shows presents no such risk (Fig 1).

Previous studies have shown that educational programs have longer scene lengths than noneducational programming. This suggests that it might in fact be the pacing of programs that may overstimulate the developing brain. Still, other plausible hypotheses for why content may matter exist.

By displacing opportunities for pretend play, television may displace opportunities for specific cognitive development that is subsequently important for executive function. It is noteworthy that among older children, *Mister Rogers’ Neighborhood* has been shown to enhance pretend play and imagination, whereas noneducational and violent programming has been shown to inhibit them.

Television generally and children’s noneducational shows specifically use language that is essentially adult-like in its pacing and pronunciation. This language, therefore, is quite different from “motherese,” the special form of language that mothers use instinctively with their young children. As spoken in live social interactions, motherese changes with the abilities of the developing infant in a way that is attuned to his or her growing ability level. Developmental neuroscientists have hypothesized that this special acoustic signal, especially as delivered in social settings, taps fundamental neurologic processes that are involved in promoting language in the developing brain of young infants. The explicitly interactive component of motherese is helpful both to give the parent time to modulate his or her performance in accordance with the needs of the child and for teaching the natural give and take of language. Noneducational television language duplicates few of the essential features of motherese. Educational shows such as *Mister Rogers’ Neighborhood*, by contrast, mimic the noninteractive elements of motherese to the extent possible. We hypothesize, therefore, that by displacing frequent interactions with adult caregivers, a heavy diet of noneducational television will delay a child’s language development.

Finally, self-regulation of affect may be impeded by noneducational television because of its frequently loud and aggressive content. Animated television shows are notoriously violent, and an assessment of animated feature films found that 100% of the G-rated animated films that were released in the United States from 1937 to 1999 contained at least some violence. Even the commercials during children’s programs are violent and conflict ridden. By contrast, prosocial content of the type found in *Mister Rogers’ Neighborhood* can promote prosocial behavior, at least among older children. We propose that the content typically found on television and in videos does not promote emotional self-regulation but rather inhibits it.

Although this theory emphasizes long-term developmental trajectories, the results presented here are also consistent with studies in the psychology literature that have evaluated the effects of television immediately after viewing. A 1973 study found that children who watched *Mister Rogers’ Neighborhood* in a laboratory setting or played instead of watching any television had greater tolerance for delay immediately afterward than children who had watched *Batman*.

In another study, those who watched *Power Rangers* had shorter attentional capacity immediately afterward than those who had watched *Mister Rogers’ Neighborhood* or played. By contrast but consistent with the results here, another study found no difference in impulsivity after children watched 40 minutes of slow-paced or fast-paced versions of *Sesame Street*.

In referring to educational programming, this discussion has mentioned *Mister Rogers’ Neighborhood* and *Sesame Street* because they have been extensively researched. Other educational programs, including some excellent ones on other networks, presumably have similar beneficial or at least nonharmful effects.

There are several limitations worth discussing. First, the outcome measure is only an approximation of psychopathology. There may be biases in the parental reporting of problematic symptoms, and there is certainly considerable measurement error. Although the effects of any potential biases are unknowable, the effect of measurement error would be to reduce the likelihood of finding an effect, a conservative bias. The outcome measure here may not be specific to ADHD but may serve as a proxy for other problems of executive function. This limitation suggests that the analysis is unable to distinguish whether the association between early noneducational television viewing and subsequent attention problems is specific to symptoms of ADHD or whether, as seems more likely, it is general to deficits of executive function. In a subanalysis of these data that used a similarly dichotomized outcome on a 3-item scale com-

![FIGURE 1](https://via.placeholder.com/150)

**FIGURE 1**

ORs of having symptoms of attentional problems according to media content.
posed of “restless,” “impulsive,” and “difficulty concentrating,” the results were virtually identical to those reported here. So were the results of a subanalysis of a 2-item scale composed of “has obsessions” and “easily confused.” We do not attach much importance to these subanalyses because they involve the use of scales that have not been psychometrically vetted, but these results are consistent with a general association with symptoms that are broader than just those of ADHD. A recent study found a contemporaneous association between television viewing and ADHD using more robust measures of the outcome. Future research should attempt to replicate the results here with a more precise and accurate measure of the developmental outcome, whether ADHD, attentional problems, or executive function. Second, although time diary data such as those used here have been found to have a high degree of validity, the measure of early exposure to television is based on a small number of days, which may have led to measurement error in the exposure. Again, this problem should be rectified in future research with more accurate assessments of early viewing but represents a conservative effect in this analysis. Finally and most fundamental, the results here are observational rather than experimental; accordingly, the results suggest only an association, not a causal relationship. The longitudinal design goes some way toward reducing the likelihood that the results are attributable to reverse causality (viz, the possibility that children with attentional problems are more likely to watch television than children without such problems), particularly if one believes that attentional problems are not fully manifest before age 3. The variables that were included to control for parenting style (early cognitive support and early emotional stimulation) go some way toward reducing the likelihood that the results are attributable to residual confounding, but these solutions are not perfect, and there remains the possibility that the results here could be partially or entirely explained by either reverse causality or residual confounding. These limitations are balanced by several strengths, including the study’s longitudinal design, the control for several attributes that are likely to be highly correlated with parenting behavior, and data on television viewing by content type.

CONCLUSIONS

Early television viewing is associated with the subsequent development of problems with attention regulation. This study shows that the association is specific to noneducational television, particularly violent fare, and that the association is specific to exposure before age 3. We believe that the associations between early television viewing and attentional problems are now sufficiently well established to warrant a randomized trial to assess causality in this association. There would be considerable benefit of such a study both if experimentally induced television reduction were associated with differences in attention-regulation, in which case parents could be appropriately advised about the impact of early television viewing on subsequent attention problems, and if such a television reduction were not associated with such differences, in which case parents could be reassured about the lack of such an impact.

ACKNOWLEDGMENTS

Dr Zimmerman’s participation was supported through National Institute of Mental Health grant 1 K01 MH06446-01A1. Throughout this project, Dr Zimmerman had full access to all of the data in the study and takes responsibility for the integrity of the data and accuracy of the data analysis.

REFERENCES


34. Barr R, Hayne H. It’s not what you know, it’s who you know: older siblings facilitate imitation during infancy. *Int J Early Years Educ*. 2003;11:7–21


39. StataCorp. *Stata Statistical Software [computer program]*. Release 9.0. College Station, TX: Stata Corp; 2005


44. Locke JL. *The Child’s Path to Spoken Language*. Cambridge, MA: Harvard University Press; 1993


ARTICLE

Violent Television Viewing During Preschool Is Associated With Antisocial Behavior During School Age

Dimitri A. Christakis, MD, MPH, Frederick J. Zimmerman, PhD

Department of Pediatrics, Child Health Institute, University of Washington, Seattle, Washington; Department of Health Services, Seattle, Washington; Seattle Children’s Research Institute, Children’s Hospital and Regional Medical Center, Seattle, Washington

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The effect of violent television programming on preschoolers’ behaviors is poorly understood. The objective of this study was to test the hypothesis that exposure to violent television viewing when children are 2 to 5 years of age would be associated with antisocial behavior at ages 7 to 10.

METHODS. Data were derived from the Panel Study of Income Dynamics. Our primary outcome was being in the 88th percentile of the Behavioral Problem Index antisocial subdomain. Our primary predictor was exposure to violent screen content.

RESULTS. Data were available for 184 boys and 146 girls at both time periods. Adjusting for baseline Behavioral Problem Index scores and age, parental education, maternal depression, and cognitive and emotional support, violent television programming was associated with an increased risk for antisocial behavior for boys but not for girls. Neither educational nor nonviolent programming was associated with increased risk for boys or girls.

CONCLUSIONS. Viewing of violent programming by preschool boys is associated with subsequent aggressive behavior. Modifying the content that is viewed by young children may be warranted.
Children who are younger than 6 years are reported to experience an average of ~2 hours of screen time each day, with approximately half of that watching television and the remainder divided between DVD/videos, computer use, and video games.2–4 Many argue that more important than how much they watch is what they watch.3–5 Unfortunately, 95% of children watch programs that are not specifically produced for young audiences6; however, even commercial television programming that is designed specifically for children can still represent a substantial risk to young children, especially with respect to violence and aggression.7 In fact, the level of violence in commercially aired programs that are intended for children exceeds the level in adult programs, including comedies, dramas, and even music videos and reality shows.7 G-rated films for children are no better: every single G-rated film that was released to theaters in the United States up to 1999 contains violence, and half show at least 1 character rejoicing in violence by cheering or laughing.8

Cross-sectional and quasi-experimental studies of television viewing among school-aged children and adolescents have found it to be associated with aggression.9–11 Experimental designs have confirmed that reducing television can reduce aggression among 9-year-olds.12,13 Considerably less attention has been given to the effects of television on preschool children.11 Although a few longitudinal studies of television viewing before age 5 have shown it to be a potential risk factor for the subsequent development of bullying and aggression measured in early elementary school,14,15 these studies did not distinguish between types of programs but rather used overall television viewing as their exposure of interest. Among media researchers, the evolving consensus is that content is a critical mediator of the effects of television on children, but data in support of this for young children are sparse.16,17 We therefore conducted a study to test the hypothesis that exposure to violent programming during the preschool period would be associated with subsequent aggressive behavior.

METHODS
We conducted a longitudinal study using data from the Panel Study of Income Dynamics (PSID). The study protocol was approved by the Children’s Hospital and Regional Medical Center Institutional Review Board.

Data Source
The PSID is a longitudinal study that was begun with 4800 families in 1968 with a variety of funding sources and overseen by the National Science Foundation. In 1997, a Child Development Supplement (CDS), which is a questionnaire administered to the primary caregivers of 3563 children aged 0 to 12, was added. The recruiting, eligibility, and attrition of the PSID and CDS have been described elsewhere.18,19

The questionnaire, funded by the National Institute of Child Health and Human Development, included detailed demographic data, psychological and behavioral assessment of parents and children, and time use diary data from 1 randomly chosen weekday and 1 randomly chosen weekend day during a school year (September through May). These diaries include primary and secondary activities during a 24-hour period. Such time diaries have been used extensively in research and have excellent validity when compared with direct observation of activities.20,21 Among eligible households in the core PSID sample, 1997 CDS data were obtained for 88%. In 2002, the respondents of the first CDS were followed up with a second, similar CDS (CDS-II). The follow-up rate from the 1997 CDS to the 2002 CDS was 91%.

Subjects
We included all children who were between the ages of 24 and 60 months in 1997 and who had follow-up data 5 years later (ie, when the were 7–9 years of age).

Outcome Variable
As part of the PSID survey, parents complete the Behavioral Problem Index (BPI) which is derived from questions from the Achenbach Child Behavior Checklist. A subdomain of the BPI assesses antisocial behavior, which has been well validated with Cronbach’s α of .66.22 The following statements are included in this subdomain: My child cheats, is mean to others, feels no regret, is destructive, is disobedient at school, and has trouble with teachers. Each statement had 3 Likert-type answers ranging from 1 for “not true” to 3 for “often true.” These raw scores were averaged and then dichotomized at 1.4, which classified 12.2% of children as having problems with antisocial behavior. This cutoff was used in an effort to approximate the 90th percentile that has been used in other studies that have used the BPI.23,24 Children at that percentile on the entire BPI have been previously found to be at risk for mental health referrals.25

Main Predictors
We used content data for programming that was derived from the time diaries. A coding system was developed previously by others to classify television shows and movies on video (hereafter both referred to as “shows”) by content in several dimensions.26 An educational attribute was assigned when the show had a clear intent to educate, with an explicit cognitive or prosocial component according to the following criteria: (1) the program teaches a lesson with content similar to that found in schools (eg, math skills, reading skills, other school readiness skills), and (2) the program teaches a lesson about appropriate behavior or interpersonal interactions (eg, sharing, friendships, drug education). Violent content was ascribed when “violence was a central and
integral part of the plot or of the main characters’ occupations, if the lead characters’ main purpose was to fight or flee from violence, or if there was more violence in the program than would be expected in the everyday life of a child.”27 Although the term “violent” is used here, it should be understood that the definition includes hostile language, threatening behavior, and cartoon violence as well as realistic violence. A minimum of 2 coders evaluated each show for violence and educational content, with a interrater \( \kappa \) score of 0.81.19,27 Differences were resolved by discussion. For this study, we classified shows into 3 categories: educational, nonviolent entertainment (ie, not violent and not educational), and violent entertainment (ie, violent and not educational). No educational shows contained violence.

Some shows could not be coded for violence, either because the name was inadequately reported (eg, “cartoons,” “Channel 13”) or because the researchers could not evaluate the violent content of an uncommon but named video or show. These shows, 20% of the total, were included in the “nonviolent entertainment” category. Examples of selected most popular shows are presented in Table 1. Television viewing was modeled as a continuous variable as number of hours on a typical day.

Covariates
We adjusted for the following covariates: race and ethnicity, gender, age at survey completion, paternal and maternal education (defined as highest grade completed at the baseline survey); paternal presence in the household; maternal depression as measured by the Center for Epidemiologic Studies Depression Scale; and emotional and cognitive stimulation at the baseline point as measured by the Home Observation for Measurement of the Environment scale,28 which has been used extensively in studies measuring household function. We also adjusted for antisocial behavior at the baseline period using the same subdomain of the BPI that we used as our outcome variable at the follow-up period. Finally, we adjusted for harsh physical punishment and parental coping in the baseline period. Physical punishment was measured as retrospective parental report of whether they spanked their child before age 2 or whether they spanked their child after age 2, with “never spanked” as the reference category. We also included the Parental Coping Problems scale, which included 5 statements:

1. [Child] seems to be harder to care for than most children.
2. There are some things that [he/she] does that really bother me a lot.
3. I find myself giving up more of my life to meet [child’s] needs than I ever expected.
4. I often feel angry with [child].
5. I would be doing better in my life without [child].

Higher scores indicate poorer coping.

Statistical Analysis
We used logistic regression to determine the independent association between types of programming at baseline and antisocial behavior at follow-up. We formally tested for effect modification of gender on content by testing an interaction term gender \( \times \) content in our model. Because the \( P \) value associated with this term was significant, we stratified our results on the basis of gender. Appropriate sampling weights were used to provide nationally representative samples. We also conducted locally weighted smoothing regression (Lowess) to plot the bivariate association between early violent television exposure and the probability of demonstrating antisocial behavior at follow-up.29,30 All analyses were conducted by using Stata 9.0 (Stata Corp, College Station, TX).

RESULTS
Data were available for 184 boys and 146 girls at both time periods, 54% of the children in age range in the

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Selected Popular Shows for Each Viewing Category According to Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category</strong></td>
<td><strong>Boys Aged 2–4 y</strong></td>
</tr>
<tr>
<td>Educational</td>
<td>Barney</td>
</tr>
<tr>
<td></td>
<td>Sesame Street</td>
</tr>
<tr>
<td></td>
<td>Magic School Bus</td>
</tr>
<tr>
<td></td>
<td>Arthur</td>
</tr>
<tr>
<td></td>
<td>Blue’s Clues</td>
</tr>
<tr>
<td>Nonviolent entertainment</td>
<td>Toy Story</td>
</tr>
<tr>
<td></td>
<td>Flintstones</td>
</tr>
<tr>
<td></td>
<td>Rugrats</td>
</tr>
<tr>
<td></td>
<td>Land Before Time</td>
</tr>
<tr>
<td></td>
<td>Basketball</td>
</tr>
<tr>
<td>Violent entertainment</td>
<td>Power Rangers</td>
</tr>
<tr>
<td></td>
<td>Football</td>
</tr>
<tr>
<td></td>
<td>Star Wars</td>
</tr>
<tr>
<td></td>
<td>Space Jam</td>
</tr>
<tr>
<td></td>
<td>Spider-Man</td>
</tr>
</tbody>
</table>
total sample during this interval. Those with missing data were most commonly missing the measurement of antisocial behavior at baseline, which was not regularly collected for those who were younger than 3 years at baseline. Accordingly, only 4% of the estimation sample was younger than 36 months, as opposed to 35% of the full sample. The mean age of study participants at the baseline period was 49 months; 44% were female. Complete demographic data are presented in Table 2.

Adjusting for all covariates concurrently, violent television programming was associated with an increased risk for antisocial behavior overall (odds ratio [OR]: 2.20; 95% confidence interval [CI]: 1.35–3.60). In the stratified analysis, however, the association was present for boys (OR: 4.10; 95% CI: 2.09–8.02) but not girls (OR: 0.39; 95% CI: 0.04–3.74). Nonviolent programming was not associated with subsequent antisocial behavior in boys (OR: 1.76; 95% CI: 0.64–4.87) or girls (OR: 1.38; 95% CI: 0.68–2.83). Educational programming also was not associated with increased risk for boys (OR: 0.41; 95% CI: 0.09–1.86) or girls (OR: 0.63; 95% CI: 0.11–3.73). Fully adjusted regression models are presented in Table 3. The adjusted results for all of the television programming types are presented in Figure 1. Removal of the single outlier in Figure 2 who watched 5 hours of violent television daily attenuated but did not eliminate the association between watching violent programming and antisocial behavior (OR: 3.27; 95% CI: 1.45–7.39).

**DISCUSSION**

We found that watching violent entertainment programming at age of 2 to 4 years was associated with significantly increased risk for antisocial behavior at ages 7 to 9 for boys but not for girls. Furthermore, we found no significant effect of other types of programming on antisocial behavior for either boys or girls during the same period. Our findings confirm and extend findings of others that exposure to violence on screen can promote aggression in real life. The differential effects of content, particularly that the point estimates for educational programming were in the direction of a protective effect, are important in that they suggest that alternative programming types could offer behavioral benefits to children without necessarily reducing overall viewing time. Others have found that select programming can promote prosocial behaviors in preschool children.31–33

It is also interesting that we found an association for boys and not for girls. This effect modification may be attributable to socialization differences between genders, genetic predispositions toward aggression, or perhaps even to the selection of programs because the shows that they watch are not identical. It is not possible in this data set to tease out the marginal contribution of each of these factors.

These results are important because aggressive behavior in the early childhood years has been repeatedly linked to violence in later youth and adolescence.34–37 This later youth violence does not materialize suddenly from nowhere but rather develops as a continuum with escalating aggression.38–42 Although minor aggressive behaviors are present from infancy, peer-directed aggression begins during the toddler and preschool years, and the onset of real physical fighting and interpersonal violence occurs later during youth and adolescence. The toddler and preschool years constitute the time during which most children learn to use nonaggressive alternatives preferentially over aggressive behavior; when that does not occur, young children can continue on a trajectory of aggression.43

Our findings must be interpreted in light of several limitations. First, the observational nature of our study precludes definitive establishment of a causal relation-

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall (N = 376)</th>
<th>Boys (n = 211)</th>
<th>Girls (n = 165)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at baseline, mean (SD), mo</td>
<td>48.72 (7.50)</td>
<td>48.98 (7.32)</td>
<td>48.51 (7.46)</td>
</tr>
<tr>
<td>Race, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>36</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>Hispanic</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Maternal education, mean (SD) y</td>
<td>12.91 (2.42)</td>
<td>13.08 (2.37)</td>
<td>12.70 (2.47)</td>
</tr>
<tr>
<td>Paternal education, mean (SD) y</td>
<td>12.61 (2.57)</td>
<td>12.65 (2.62)</td>
<td>12.55 (2.49)</td>
</tr>
<tr>
<td>Father present in home, %</td>
<td>74</td>
<td>77</td>
<td>70</td>
</tr>
<tr>
<td>Spanked before age 2, %</td>
<td>27</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>Spanked after age 2, %</td>
<td>60</td>
<td>64</td>
<td>55</td>
</tr>
<tr>
<td>Depression score (mother), mean (SD)</td>
<td>9.44 (3.98)</td>
<td>9.39 (3.82)</td>
<td>9.50 (3.80)</td>
</tr>
<tr>
<td>Cognitive stimulation score, mean (SD)</td>
<td>10.89 (2.06)</td>
<td>10.86 (2.10)</td>
<td>10.94 (2.01)</td>
</tr>
<tr>
<td>Emotional support score, mean (SD)</td>
<td>8.22 (1.34)</td>
<td>8.24 (1.38)</td>
<td>8.22 (1.35)</td>
</tr>
<tr>
<td>Coping problems score, mean (SD)</td>
<td>8.57 (3.51)</td>
<td>8.66 (3.62)</td>
<td>8.19 (3.33)</td>
</tr>
<tr>
<td>Television, mean (SD) h/d</td>
<td>Educational 0.42 (0.60) 0.47 (0.60) 0.38 (0.59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonviolent entertainment 0.91 (1.00) 0.83 (0.94) 1.00 (1.08)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Violent entertainment 0.54 (0.80) 0.55 (0.82) 0.51 (0.78)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
However, that we were able to control for antisocial behavior at the baseline period in our regression analyses and that the hypothesis tested is based on experimentally confirmed data in older children lend support to the possibility that in fact these associations are causal. Second, we used time diaries to determine shows viewed. Although imperfect, such diaries have been shown to be valid.\textsuperscript{44} Third, our cutoff for antisocial problems at the 88th percentile is somewhat arbitrary; however, it is consistent with the 90th percentile that others have used. We were lacking data on some particular shows when parents merely reported that their child was watching cartoons. Our procedure to classify all such shows as nonviolent entertainment might bias our findings toward findings of aggression effects for such programming, because a great deal of cartoons are in fact violent; however, by not classifying them as violent, we should not have biased our estimate for violent programming. Despite these limitations, our results suggest that modification of the media diet during the preschool years may have long-term effects on aggressive behavior in children.

### TABLE 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Boys, OR (95% CI)</th>
<th>Girls, OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal education</td>
<td>0.63 (0.47–0.86)</td>
<td>0.47 (0.30–0.75)</td>
</tr>
<tr>
<td>Emotional support</td>
<td>1.24 (0.63–2.46)</td>
<td>0.51 (0.29–0.89)</td>
</tr>
<tr>
<td>Cognitive stimulation</td>
<td>0.64 (0.44–0.94)</td>
<td>1.26 (0.86–1.86)</td>
</tr>
<tr>
<td>Maternal depression</td>
<td>1.68 (1.27–2.22)</td>
<td>1.23 (1.03–1.47)</td>
</tr>
<tr>
<td>Father present in household</td>
<td>0.67 (0.24–1.87)</td>
<td>10.29 (0.45–236.00)</td>
</tr>
<tr>
<td>Spanked before age 2</td>
<td>7.65 (0.73–80.21)</td>
<td>21.24 (0.65–695.00)</td>
</tr>
<tr>
<td>Spanked after age 2</td>
<td>0.67 (0.09–5.07)</td>
<td>21.24 (0.65–695.75)</td>
</tr>
<tr>
<td>Parental coping problems scale</td>
<td>0.99 (0.86–1.14)</td>
<td>1.23 (1.01–1.50)</td>
</tr>
<tr>
<td>BPI high at baseline</td>
<td>1.69 (0.46–6.25)</td>
<td>1.81 (0.23–14.18)</td>
</tr>
<tr>
<td>Television viewing at baseline, h/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational</td>
<td>0.45 (0.12–1.78)</td>
<td>0.85 (0.25–2.88)</td>
</tr>
<tr>
<td>Nonviolent entertainment</td>
<td>2.82 (0.86–9.24)</td>
<td>1.38 (0.71–2.65)</td>
</tr>
<tr>
<td>Violent entertainment</td>
<td>4.10 (2.09–8.02)</td>
<td>0.39 (0.04–3.74)</td>
</tr>
</tbody>
</table>

Also adjusted for race/ethnicity and age at interview.
FIGURE 2
Lowess graphs of antisocial problems regressed on violent content. A, Boys; B, girls.

REFERENCES
33. Silverman LT, Sprafkin JN. The effects of Sesame Street's
prosocial spots on cooperative play between young children. 
*J Broadcast.* 1980;24:135–147

34. Brame B, Nagin DS, Tremblay RE. Developmental trajectories of physical aggression from school entry to late adolescence. 


36. Ladd GW, Burgess KB. Charting the relationship trajectories of aggressive, withdrawn, and aggressive/withdrawn children during early grade school. 
*Child Dev.* 1999;70:910–929

37. Nagin DS, Tremblay RE. Parental and early childhood predictors of persistent physical aggression in boys from kindergarten to high school. 
*Arch Gen Psychiatry.* 2001;58:389–394

*Dev Psychol.* 2003;39:222–245


40. Fraser MW. Aggressive behavior in childhood and early adolescence: an ecological-developmental perspective on youth violence. 

41. Loeber R, Hay D. Key issues in the development of aggression and violence from childhood to early adulthood. 

42. Nagin D, Tremblay RE. Trajectories of boys’ physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. 
*Child Dev.* 1999;70:1181–1196

*Pediatrics.* 2004;114(1). Available at: www.pediatrics.org/cgi/content/full/114/1/e43


EARLY THIMEROSAL EXPOSURE AND NEUROPSYCHOLOGICAL OUTCOMES AT 7 TO 10 YEARS

“Background. It has been hypothesized that early exposure to thimerosal, a mercury-containing preservative used in vaccines and immune globulin preparations, is associated with neuropsychological deficits in children. 

Methods. We enrolled 1047 children between the ages of 7 and 10 years and administered standardized tests assessing 42 neuropsychological outcomes. (We did not assess autism-spectrum disorders.) Exposure to mercury from thimerosal was determined from computerized immunization records, medical records, personal immunization records, and parent interviews. Information on potential confounding factors was obtained from the interviews and medical charts. We assessed the association between current neuropsychological performance and exposure to mercury during the prenatal period, the neonatal period (birth to 28 days), and the first 7 months of life. 

Results. Among the 42 neuropsychological outcomes, we detected only a few significant associations with exposure to mercury from thimerosal. The detected associations were small and almost equally divided between positive and negative effects. Higher prenatal exposure was associated with better performance on one measure of language and poorer performance on one measure of attention and executive functioning. Increasing levels of mercury exposure from birth to 7 months were associated with better performance on one measure of fine motor coordination and on one measure of attention and executive functioning. Increasing mercury exposure from birth to 28 days was associated with poorer performance on one measure of speech articulation and better performance on one measure of fine motor coordination. 

Conclusions. Our study does not support a causal association between early exposure to mercury from thimerosal-containing vaccines and immune globulins and deficits in neuropsychological functioning at the age of 7 to 10 years.”


Noted by JFL
Healing of Nonhymenal Genital Injuries in Prepubertal and Adolescent Girls: A Descriptive Study

John McCann, MD, Sheridan Miyamoto, MSN, FNP, Cathy Boyle, MSN, PNP, Kristen Rogers, PhD

Children’s and Adolescent’s Research and Evaluation Center (CAARE) Diagnostic and Treatment Center, University of California, Davis, Children’s Hospital, Sacramento, California

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The objective of this study was to identify the healing process and outcome of nonhymenal injuries in prepubertal and pubertal girls.

METHODS. This multicenter, retrospective project used photographs to document the healing process and outcome of nonhymenal genital injuries in 239 prepubertal and pubertal girls whose ages ranged from 4 months to 18 years.

RESULTS. The genital injuries sustained by the 113 prepubertal girls consisted of 21 accidental or noninflicted injuries, 73 injuries secondary to abuse, and 19 injuries of unknown cause. All 126 pubertal girls were sexual assault victims. These nonhymenal genital injuries healed at various rates depending on the type and severity. There was no statistical difference in the rate of healing between the 2 groups. Abrasions disappeared by the third day after injury. Edema was no longer present by the fifth day. Ecchymosis (bruising) resolved within 2 to 18 days depending on the severity. One prepubertal girl still had a labial hematoma at 2 weeks. Submucosal hemorrhages of the vestibule and fossa navicularis resolved between 2 days and 2 weeks. Petechiae and blood blisters proved useful for approximating the age of an injury. Petechiae were gone by 24 hours, whereas blood blisters were detected at 30 days in a prepubertal girl and 24 days in a pubertal girl. The depth of a laceration determined the time required for it to heal. Superficial vestibular lacerations seemed healed in 2 days, whereas deep perineal lacerations required up to 20 days. The appearance of new blood vessel formation was detected only in prepubertal girls, whereas scar tissue formation occurred only after a deep laceration in both groups.

CONCLUSIONS. The majority of these nonhymenal genital injuries healed with little or no evidence of previous trauma. The time required for resolution varied by type, location, and severity.
The evaluation of a female child or adolescent who is suspected of having been sexually abused requires a careful evaluation. It begins with a detailed history of the event(s) followed by a complete physical examination and, if appropriate, the collection of forensic material. Although the interpretation of the outcome of the findings has received a good deal of attention, there are still many unanswered questions.\textsuperscript{1,12}

There is relatively little information about the healing process of a female’s genital injury. Except for the report by Heppenstall-Heger et al,\textsuperscript{10} most studies have focused on the outcome of hymenal injuries in prepubertal girls.\textsuperscript{11,12} Very little is known about the amount of time that it takes a genital injury to resolve. It is unclear whether the type of injury or its location or the developmental stage of the girl makes a difference. It is unknown whether there is a sequence or a time frame in this process that could be used to determine the age of a healing injury. It is also unknown whether there are specific findings that could help to identify the approximate age of an injury. This study, along with its companion report\textsuperscript{13} on the healing of hymenal injuries, was designed to help answer these questions.

\section*{METHODS}

\subsection*{Recruitment}

The patients for this multicenter study were recruited from medical centers throughout the United States. The majority of these cases were obtained through the use of the Heller Society’s Listserv. The members of this honorary society are recognized for their expertise in the evaluation and treatment of abused and neglected children and adolescents. The participants were asked to provide pertinent medical information and photographs of any female child or adolescent who had sustained a genital injury from any cause. Patients from birth through 17 years of age were eligible. In addition to the photographic documentation of an injury, all patients were required to have at least 1 follow-up examination. Because this was a retrospective, convenience-sample study, the period between an injury and the follow-up examination was not uniform. Each center’s institutional review board authorized its center’s participation in the project.

\subsection*{Historical Information}

The participants provided the authors with a summary of the portion of a patient’s medical chart that pertained to the genital injury. The information requested included the individual’s birth date and ethnicity, the examiner’s opinion as to the cause of the injury, and the examination method used. The time and date of all examinations became part of a computer-generated database. The patient’s computerized medical chart and photographs were assigned a number to protect the individual’s identity.

\subsection*{Photographic Documentation}

Photographic documentation by the participating institutions was captured through the use of a variety of recording methods. The most common recording device was either a 35-mm camera with a macro lens or a camera that was mounted on a colposcope. Several centers provided images that had been captured through the use of digital still or video cameras. Prints of the images were provided by each center.

\subsection*{Analysis of the Photographs}

The patients were examined by a variety of methods. These included the supine labial separation method, the supine labial traction technique, and the prone knee-chest position. When a patient had been examined by >1 method, we divided the photographs into separate envelopes on the basis of the method used. Each photograph was evaluated in the presence of all 3 medical authors. During the evaluation, the authors were blinded to the cause of the patient’s injury and to the contributor’s interpretation of the findings. An agreement by all 3 medical authors was required before the interpretation of a finding was recorded on a worksheet and entered into the database.

\subsection*{Analysis of the Patients}

We divided the patients into 2 groups on the basis of the hormonal effect on the hymen. The first group consisted of the girls whose hymen showed no estrogen effect. Their hymens tended to be thin, delicate membranes with relatively smooth edges. The few girls who were younger than 3 years and retained some visual evidence of endogenous estrogen were placed in the first group that is referred to as prepubertal girls. The second group consisted of the older girls whose hymen did show an estrogen effect. Their hymens tended to be thicker and more redundant and frequently had scalloped edges. This second group is referred to as pubertal girls.

\subsection*{Types of Nonhymenal Genital Injuries}

We subdivided the nonhymenal injuries into abrasions, contusions, and lacerations (see Appendix). Evidence of a contusion included the presence of blood blisters, edema, hematomas, petechiae, and submucosal hemorrhages. The abrasions and contusions were classified further as to their size and color. The severity of each abrasion and contusion was categorized through the use of the term mild, moderate, or marked. Each of these characteristics was compared with those noted during the initial examination, and the time between examinations was recorded.

The depth of a nonhymenal genital laceration was based on its appearance. A laceration was recorded as
superficial when it did not extend into the subcutaneous tissue, intermediate when subcutaneous tissue could be visualized, and deep when the injury penetrated into the underlying muscle. Completion of the healing process was defined by the disappearance of the signs of an acute injury and the cessation of changes in the depth and configuration of a laceration. An acute injury was defined as the presence of a fresh-cut surface and/or evidence of active bleeding. New blood vessel (neovascularity) and scar tissue formation along with the appearance of the finding at the time of the last evaluation was also documented.

**Interobserver Reliability**
Individually, we performed a blinded reexamination of a random sample of 10% (n = 25) of the cases to assess and measure the reliability of the original agreement on the interpretation of a finding. $\kappa$ statistics were used to determine this interobserver reliability. The $\kappa$ scores ranged from 0.46 to 1.0 (moderate to excellent). On the basis of the interpretation of $\kappa$ results by Landis and Koch\(^14\) as well as Fleiss,\(^15\) the authors concluded that the results from this study are sufficiently reliable.

**Statistical Analysis**
We entered all collected data into an Access (Microsoft, Redmond, WA) database. We documented the photographic findings on a worksheet and systematically entered it into the created database. These data were then transferred directly into SPSS (SPSS, Inc, Chicago, IL) for analysis. Descriptive statistics were used to show the results of each of the 3 examination methods. Data were analyzed by using $t$ tests, $\chi^2$, Yates continuity correction or Fisher’s exact tests, and Mann-Whitney $U$ tests, when appropriate. Statistical significance was defined as $P < .05$.

**RESULTS**
The patients consisted of 239 female children and adolescents from 4 months to 18 years of age. There were 113 (47%) prepubertal girls and 126 (53%) pubertal girls. Of the 113 prepubertal children, 50% were white, 16% were black, 26% were Hispanic, 4% were Asian, and 4% were of a mixed race. Of the 126 pubertal girls, 48% were white, 28% were black, 14% were Hispanic, 4% were Asian, and 6% were of a mixed race.

**Timing of Examinations and the Cause of Injuries**
The period between an injury and the initial examination ranged from 1 hour to 3 days. A total of 164 (69%) of the 239 girls were seen within 24 hours of their injury, 208 (87%) were examined within 48 hours, and the other 31 (13%) were first evaluated between 48 and 72 hours after their injury. The mean time between an injury and the first examination was 24 hours. The longest a prepubertal girl was followed was 2.6 years. The average follow-up period was 9.9 months. For the 126 pubertal girls, the soonest reevaluation also occurred within 24 hours of the assault. The longest a pubertal girl was followed was 3.7 months. The average follow-up period was 61 days. The causes of the injuries in the 113 prepubertal girls, as determined by the contributors, included 21 (19%) accidental or noninflicted injuries, 73 (65%) injuries secondary to abuse, and 19 (17%) injuries of unknown cause. All 126 pubertal adolescents were said to be victims of a sexual assault.

**Summary of the Findings**
Because of the nature of this study, the timing of both the initial and the follow-up examinations varied as a result of the circumstances of each case. During the follow-up examination, the number of days since the injury and the status of each nonhymenal genital abrasion, contusion, or laceration were recorded.

Tables 1 and 2 are a compilation of the period required for a nonhymenal abrasion or a contusion to resolve. The healing process was recorded as follows:

- “Last detected” identifies the day on which a finding was last documented in patients with a particular injury.
- “Earliest disappearance” identifies the day on which a particular finding was no longer identified in any 1 patient.
- “Gone” identifies the first examination day on which a finding was no longer seen in any of the patients with that finding.
- “Never seen” represents a finding that was never identified during a follow-up examination in any of the patients with that type of injury. Unfortunately, in this case, there was no way of knowing when such an injury had actually disappeared.

For example, in Table 1, of the 21 (19%) prepubertal girls with a labial abrasion, for whom there were a total of 31 follow-up examinations, the last time this finding was seen on a follow-up examination in any of the girls was on the fourth day after an injury. The earliest day on which a labial abrasion was not detected was on the third day in 1 of 9 patients with this finding. No labial abrasions were identified during any of the other follow-up examinations (20 of 20) beginning with the fifth day after injury. Eight other prepubertal girls with this type of injury were reexamined during the next 5 days, and none of them had evidence of an abrasion on the labia.

The abrasions in the pubertal girls as found in Table 1 are used as an example of the term “never seen.” Nineteen (15%) pubertal girls had what initially seemed to be a labial abrasion. Of the 19 follow-up evaluations, no abrasions were ever detected (“never seen”) during a follow-up evaluation (19 of 19). The soonest that any of
these girls (1 of 1) were reexamined was on the third day after their assault. During the next 7 days, 11 more of the pubertal girls with this finding were reevaluated and none had signs of their previous injury.

Nonhymenal Genital Contusions

**Blood Blister**
A thin blood-filled vesicle (blood blister) was found in the vestibule of 2 prepubertal and in 6 pubertal girls (Table 2). In both groups, the blood blisters were usually detected during a follow-up examination as the tissue swelling receded and the extravasation of blood resolved. The earliest identification occurred on the fifth day after injury in a prepubertal girl and on the sixth day after assault in a pubertal girl. Once formed, these 2- to 5-mm blood-filled vesicles shrank in size as they healed.

One of the 2 prepubertal girls with this finding still had a blood blister 30 days after her injury (see Table 2). It is unknown when this blister eventually disappeared. The blood blister in the other prepubertal girl was gone by the time she returned 48 days after her injury.

Blood blisters were present in 6 pubertal girls on their visits 6, 18, 22, and 24 days after assault. Blood blisters were gone in 2 pubertal girls by days 30 and 33. It is unknown when the other blood blisters eventually disappeared.

**Ecchymosis**
Flat hemorrhagic lesions (ecchymosis/bruise) were detected on the keratinized tissues of the labia, perineum, and posterior fourchette (Table 2). In the prepubertal girls, straddle-type injuries accounted for 37 (67%) of the labial injuries and 13 (38%) of the posterior fourchette injuries. The earliest disappearance of a bruise in the prepubertal girls ranged from 2 to 5 days. During the first week after injury, evidence of a bruise had disappeared in 21 of 30 prepubertal girls with this finding.

Very few bruises (ecchymosis) were identified on any part of the pubertal girls’ genitalia (Table 2). Evidence of a bruise was never seen during a follow-up examination in the majority of these girls (see Table 2).

**Edema**
No sign of edema was found during a follow-up examination in any of the 10 prepubertal girls with this finding (Table 2). Edema was last detected on the fifth day in 1 of the 4 adolescent girls with signs of swelling of the tissues.

**Erythema**
The redness of tissues created by capillary congestion (erythema) constitutes a nonspecific finding; therefore, erythema is not included as a variable in this section because of its uncertain clinical significance.

---

### TABLE 1
**Healing of Nonhymenal Genital Injuries in Prepubertal and Pubertal Girls: Abrasions**

<table>
<thead>
<tr>
<th>Location of Injury</th>
<th>Group, n (%)</th>
<th>Last Detected, d (n/N)</th>
<th>Earliest Disappearance, d (n/N)</th>
<th>Gone, d (n/N)</th>
<th>Never Seen, n/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>21 (19)</td>
<td>4 (1/2)</td>
<td>3 (1/9)</td>
<td>5 (20/20)</td>
<td>—</td>
</tr>
<tr>
<td>F/U</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pub</td>
<td>19 (15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F/U</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vestibule</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>14 (12)</td>
<td></td>
<td>3 (1/1)</td>
<td>—</td>
<td>X (15/15)</td>
</tr>
<tr>
<td>F/U</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pub</td>
<td>12 (10)</td>
<td>3 (1/2)</td>
<td>5 (1/2)</td>
<td>5 (9/9)</td>
<td>—</td>
</tr>
<tr>
<td>F/U</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fossa navicularis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>5 (4)</td>
<td>7 (1/3)</td>
<td>2 (2/2)</td>
<td>7 (3/3)</td>
<td>—</td>
</tr>
<tr>
<td>F/U</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pub</td>
<td>24 (19)</td>
<td></td>
<td>3 (2/2)</td>
<td>—</td>
<td>X (26/26)</td>
</tr>
<tr>
<td>F/U</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior fourchette</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>4 (4)</td>
<td></td>
<td>2 (1/1)</td>
<td>—</td>
<td>X (6/6)</td>
</tr>
<tr>
<td>F/U</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pub</td>
<td>5 (4)</td>
<td></td>
<td>3 (1/1)</td>
<td>—</td>
<td>X (6/6)</td>
</tr>
<tr>
<td>F/U</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perineal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>4 (4)</td>
<td>17 (1/3)</td>
<td>5 (2/2)</td>
<td>20 (1/1)</td>
<td>—</td>
</tr>
<tr>
<td>F/U</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pub</td>
<td>3 (2)</td>
<td></td>
<td>6 (1/1)</td>
<td>—</td>
<td>X (3/3)</td>
</tr>
<tr>
<td>F/U</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pre indicates prepubertal girls (nonestrogenized hymen); Pub, pubertal girls (estrogenized hymen); F/U, follow-up examinations; —, no findings in this category; X, never seen.
<table>
<thead>
<tr>
<th>Type of Injury</th>
<th>Group, n (%)</th>
<th>Last Detected, d (n/N)</th>
<th>Earliest Disappearance, d (n/N)</th>
<th>Gone, d (n/N)</th>
<th>Never Seen, n/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood blister</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vestibule</td>
<td>Pre 2 (2)</td>
<td>30 (1/1)</td>
<td>48 (1/2)</td>
<td>Unknown (1/1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 6 (5)</td>
<td>24 (1/2)</td>
<td>30 (2/2)</td>
<td>33 (2/2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ecchymosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labia</td>
<td>Pre 37 (33)</td>
<td>18 (1/40)</td>
<td>2 (3/4)</td>
<td>23 (13/13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 2 (2)</td>
<td>3 (1/1)</td>
<td>10 (1/1)</td>
<td>10 (2/2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior fourchette</td>
<td>Pre 13 (12)</td>
<td>15 (1/10)</td>
<td>4 (1/2)</td>
<td>18 (5/5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 3 (4)</td>
<td></td>
<td>3 (1/3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 4</td>
<td></td>
<td></td>
<td>X (4/4)</td>
<td></td>
</tr>
<tr>
<td>Edema</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labia/perineal/posterior fourchette</td>
<td>Pre 10 (9)</td>
<td></td>
<td>5 (3/3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 14</td>
<td></td>
<td></td>
<td></td>
<td>X (14/14)</td>
</tr>
<tr>
<td></td>
<td>Pub 4 (3)</td>
<td>5 (1/1)</td>
<td>5 (1/1)</td>
<td>6 (4/4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hematoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labia</td>
<td>Pre 2 (2)</td>
<td>14 (1/1)</td>
<td>10 (1/1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 2 (2)</td>
<td></td>
<td>10 (2/2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 2</td>
<td></td>
<td></td>
<td>X (2/2)</td>
<td></td>
</tr>
<tr>
<td>Vestibule</td>
<td>Pre 1 (1)</td>
<td></td>
<td>5 (1/1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 2</td>
<td></td>
<td></td>
<td></td>
<td>X (2/2)</td>
</tr>
<tr>
<td></td>
<td>Pub 0 (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fossa navicularis</td>
<td>Pre 4 (3)</td>
<td></td>
<td>2 (2/2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 8</td>
<td></td>
<td></td>
<td></td>
<td>X (8/8)</td>
</tr>
<tr>
<td></td>
<td>Pub 0 (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Petechiae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labia</td>
<td>Pre 15 (12)</td>
<td></td>
<td>2 (4/4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 19</td>
<td></td>
<td></td>
<td></td>
<td>X (19/19)</td>
</tr>
<tr>
<td></td>
<td>Pub 2 (2)</td>
<td></td>
<td>6 (1/1)</td>
<td></td>
<td>X (2/2)</td>
</tr>
<tr>
<td></td>
<td>F/U 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vestibule</td>
<td>Pre 38 (34)</td>
<td>1 (2/3)</td>
<td>2 (2/2)</td>
<td>2 (42/42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 16 (13)</td>
<td></td>
<td>2 (1/1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 17</td>
<td></td>
<td></td>
<td>X (17/17)</td>
<td></td>
</tr>
<tr>
<td>Fossa navicularis</td>
<td>Pre 13 (12)</td>
<td>1 (1/1)</td>
<td>2 (2/2)</td>
<td>2 (17/17)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 12 (10)</td>
<td></td>
<td>2 (1/1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F/U 18</td>
<td></td>
<td></td>
<td>X (18/18)</td>
<td></td>
</tr>
<tr>
<td>Posterior fourchette</td>
<td>Pre 6 (5)</td>
<td></td>
<td>2 (1/1)</td>
<td></td>
<td>X (9/9)</td>
</tr>
<tr>
<td></td>
<td>F/U 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 2 (2)</td>
<td></td>
<td>6 (1/1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hematomas

Hematomas of the Labia
A well-defined, localized collection of blood (hematoma) was an uncommon finding in both groups of girls (Table 2). The labial hematoma in 1 prepubertal girl was gone by 10 days (see Table 2), whereas a hematoma in the other girl was still present at her 2-week follow-up examination.

Only 2 pubertal girls had hematomas on their labia. The hematomas in both girls were never seen at the time of their first reevaluation, which took place on their 10th day after assault (see Table 2).

Hematomas of the Vestibule and Fossa Navicularis
Only 1 prepubertal girl had what initially seemed to be a vestibular hematoma (Table 2). By the first follow-up examination on day 5, this well-defined, localized collection of blood (hematoma) had disseminated into the surrounding tissues. The “hematomas” of the fossa navicularis of 4 prepubertal girls evolved into submucosal hemorrhages by the second day after injury. No hematomas were found in the vestibule or the fossa navicularis of the pubertal girls.

Petechiae
Seventy-four (65%) of the 113 prepubertal girls had pinpoint, nonraised, perfectly round, purplish red spots (petechiae) on their nonhymenal genital tissues at the time of their initial examination (Table 2). No petechiae were identified on a follow-up examination in any of the pubertal girls (see Table 2). Five of the girls were reexamined between 48 and 72 hours after their assault. None had any sign of a petechial lesion.

Submucosal Hemorrhages
Submucosal hemorrhages were among the most common lesions found in the perihymenal region of the vestibule of the prepubertal girls (Table 2). In the prepubertal girls, evidence of bleeding into the areolar tissue beneath the mucosal membrane of the vestibule was found in 62 (55%) and in the fossa navicularis in 32 (28%). During the 80 follow-up examinations, the earliest disappearance of a mild vestibular submucosal hemorrhage occurred on day 2 (3 of 7). The last marked hemorrhage of the vestibule was detected on day 8 (1 of 32). One mild fossa navicularis submucosal hemorrhage had resolved (earliest disappearance) by day 3 (1 of 4). A marked submucosal hemorrhage in this same area was last detected on day 12 (1 of 20) in 1 prepubertal girl.

Fourteen (11%) of the pubertal girls were found to have submucosal hemorrhages of the perihymenal tissues in their vestibule, whereas 13 (10%) had this same type of injury in their fossa navicularis (Table 2). The earliest disappearance of a vestibular submucosal hemorrhage in a pubertal girl occurred on the seventh day after assault (1 of 2). The earliest recorded disappearance of a mild fossa navicularis hemorrhage was on day 2 (1 of 1). As the submucosal hemorrhages resolved, they became smaller in size and paler in color.

Nonhymenal Lacerations
The lacerations were classified as superficial, intermediate, or deep (Table 3). The 157 nonhymenal lacerations identified in the 113 prepubertal girls were reevaluated...
251 times. The number of the reevaluations varied according to the circumstances. Some girls were reexamined 1 time, whereas others were reexamined several times. One prepubertal girl who sustained a laceration was followed for 2.6 years. The 87 nonhymenal lacerations detected in the 126 pubertal girls were reexamined 94 times. The earliest reevaluation occurred on the third day after assault, whereas the last assessment took place 5 months after the adolescent’s assault. Despite the relatively small numbers for any given type of laceration, there seemed to be a high correlation between our assessment of depth from a photograph and the time required for the resolution of a laceration.

Labial Lacerations
Forty-seven (37%) prepubertal girls had lacerations of the labia minora or the labia majora (Table 3). Seventy-four percent of the labial tears were at the junction of these 2 structures. The deeper lacerations took on a shallower appearance as they healed. Depending on the depth, the time required for a laceration to resolve varied from 2 days to 14 days.

All 9 (7%) of the pubertal girls’ labial lacerations were superficial (Table 3). The soonest reevaluation of any of these girls occurred on the seventh day after assault (1 of 1). There was no sign of a laceration in any of these pubertal girls who were reexamined between the seventh and the 30th days after their assault.

Vestibular Lacerations
Vestibular lacerations were documented in 10 (9%) prepubertal girls (Table 3). The earliest disappearance of a laceration occurred at 2 days in 1 of 2 girls. The last time a superficial laceration was detected was on the eighth day after injury (1 of 1). Depending on the depth, evidence of the other vestibular lacerations disappeared between 6 and 11 days. The deeper lacerations took on a more superficial appearance as they healed.

None of the 6 pubertal girls with a vestibular laceration had evidence of their injury at the time of their reexamination. The earliest reevaluations occurred on the third, fifth, and 10th days after assault.
**Fossa Navicularis Lacerations**

Forty-two (37%) prepubertal girls had lacerations of their fossa navicularis (Table 3). As with the other lacerations, the deeper tears evolved into shallower-appearing injuries as they healed. The time of resolution for the prepubertal girls’ fossa navicularis lacerations ranged from 2 days for a superficial laceration to 21 days for a deep tear (17 of 17). During the follow-up examinations, 6 of 19 fossa navicularis lacerations were noted to have resolved by 7 days and 14 of 42 by 14 days.

Fossa navicularis lacerations were detected in 38 (30%) pubertal girls (Table 3). The earliest disappearance of a superficial laceration occurred on day 4 (1 of 1), whereas a deep laceration was last detected on day 10 (1 of 3). By day 13, the remaining 6 deep lacerations had healed. Both the intermediate and deep lacerations evolved into shallower-appearing tears before healing completely. On the follow-up examinations, 6 of 9 fossa navicularis lacerations had resolved by 7 days and 18 of 23 by 14 days.

**Posterior Fourchette Lacerations**

Posterior fourchette lacerations were found in 34 (27%) prepubertal girls (Table 3). Depending on the depth, evidence of the prepubertal girls’ lacerations disappeared between 3 (1 and 1) and 20 days (19 of 19). On the follow-up examinations, 7 of 18 posterior fourchette lacerations had resolved by 7 days and 23 of 47 by 14 days.

Twenty-seven (21%) pubertal girls sustained posterior fourchette lacerations (Table 3). Although the earliest disappearance of a superficial laceration was documented on the third day after assault, none of the lacerations of this depth was ever seen again (14 of 14). The same result was found with the 9 intermediate depth lacerations. The earliest disappearance of this depth laceration was recorded on the sixth day after an assault. Of the 5 deep tears, the earliest disappearance occurred on day 7 (1 of 1) and the last 1 was detected on day 10 (1 of 1). On the follow-up examinations, 5 of 5 posterior fourchette lacerations had resolved by 7 days and 12 of 13 by 14 days.

**Perineal Lacerations**

Twenty-four (21%) prepubertal girls had perineal lacerations (Table 3). Evidence of both the superficial (6 of 6) and the intermediate (5 of 5) depth lacerations were no longer visible at the time of a follow-up evaluation. However, the earliest reevaluation of a girl with a superficial perineal laceration occurred on day 12 (1 of 1). The earliest reevaluation of a prepubertal girl with an intermediate depth laceration occurred on day 3 (1 of 1). Signs of a deep laceration were still present in 1 girl at 20 days (1 of 13). On the follow-up examinations, 5 of 5 of the deep perineal lacerations had resolved by 7 days and 7 of 16 by 14 days.

The pubertal girls had no sign of any of their perineal lacerations at the time of their reevaluations (Table 3). This included 1 of 1 reevaluation of a deep laceration on the seventh day after assault. On the follow-up examinations of the deep lacerations, 1 of 1 perineal laceration had resolved by 7 days and 4 of 4 by 14 days.

**Neovascular and Scar Tissue Formation**

New blood vessel formation (neovascular changes) and scar tissue formation (thickened pale white lines) were considered to be present only when they were detected ≥30 days after an injury. Neovascular formation was found only in the prepubertal girls and then relatively infrequently (Table 4). This new blood vessel formation occurred mainly on the mucosal surfaces of the vestibule (2 of 6) and the fossa navicularis (2 of 10).

Scar tissue formation was also an infrequent finding and then detected only in the girls who had sustained a deep laceration. Scars were found more commonly in the posterior fourchette (7 of 14) and on the perineum (6 of 8) of the prepubertal girls (Table 4).
<table>
<thead>
<tr>
<th>Location of Injury</th>
<th>Group, n (%)</th>
<th>Severity, n (%)</th>
<th>Last Detected, d (n/N)</th>
<th>Earliest Disappearance, d (n/N)</th>
<th>Healed, d (n/N)</th>
<th>Never Seen, n/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labia minora/majora</td>
<td>Pre 47 (37)</td>
<td>Superficial: 28 (48)</td>
<td>6 (1/1)</td>
<td>3 (1/2)</td>
<td>7 (33/33)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 15 (36)</td>
<td>8 (1/8)^a</td>
<td>2 (1/3)</td>
<td>9 (16/16)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 4 (17)</td>
<td>13 (1/3)^a</td>
<td>3 (1/3)</td>
<td>14 (1/1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 9 (4)</td>
<td>Superficial: 9 (100)</td>
<td></td>
<td>7 (1/1)</td>
<td></td>
<td>X (9/9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 0 (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 0 (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vestibule</td>
<td>Pre 10 (9)</td>
<td>Superficial: 8 (80)</td>
<td>8 (1/1)</td>
<td>2 (1/2)</td>
<td>8 (2/2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 1 (10)</td>
<td></td>
<td>6 (2/2)</td>
<td></td>
<td>X (4/4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 1 (10)</td>
<td>11 (2/2)^a</td>
<td>8 (1/1)</td>
<td>11 (1/1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 6 (9)</td>
<td>Superficial: 6 (100)</td>
<td></td>
<td>3 (1/1)</td>
<td></td>
<td>X (8/8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 0 (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 0 (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fossa navicularis</td>
<td>Pre 42 (37)</td>
<td>Superficial: 6 (14)</td>
<td>8 (1/4)</td>
<td>2 (1/2)</td>
<td>9 (5/5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 4 (10)</td>
<td></td>
<td>4 (1/1)</td>
<td>Unknown (3/3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 32 (76)</td>
<td>14 (12/28)^a</td>
<td>7 (2/13)</td>
<td>21 (17/17)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 38 (30)</td>
<td>Superficial: 16 (42)</td>
<td>10 (1/3)</td>
<td>4 (1/1)</td>
<td>12 (12/12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 13 (34)</td>
<td></td>
<td>6 (1/2)</td>
<td>8 (9/9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 9 (24)</td>
<td>10 (1/3)^a</td>
<td>3 (1/2)</td>
<td>13 (6/6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior fourchette</td>
<td>Pre 34 (27)</td>
<td>Superficial: 8 (24)</td>
<td>13 (1/3)</td>
<td>3 (1/1)</td>
<td>14 (7/7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 7 (21)</td>
<td></td>
<td>7 (1/2)</td>
<td>18 (6/6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 19 (56)</td>
<td>18 (1/8)</td>
<td>6 (1/7)</td>
<td>20 (19/19)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 27 (21)</td>
<td>Superficial: 13 (48)</td>
<td></td>
<td>3 (1/1)</td>
<td></td>
<td>X (14/14)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 9 (33)</td>
<td></td>
<td>6 (1/1)</td>
<td></td>
<td>X (9/9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 5 (19)</td>
<td>10 (1/1)^a</td>
<td>7 (1/1)</td>
<td>13 (4/4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perineal</td>
<td>Pre 24 (21)</td>
<td>Superficial: 5 (21)</td>
<td></td>
<td>12 (1/1)</td>
<td></td>
<td>X (6/6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 3 (13)</td>
<td></td>
<td>3 (1/1)</td>
<td></td>
<td>X (5/5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 16 (67)</td>
<td>20 (1/13)^a</td>
<td>3 (1/5)</td>
<td>21 (9/9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub 7 (6)</td>
<td>Superficial: 1 (14)</td>
<td></td>
<td>18 (1/1)</td>
<td></td>
<td>X (1/1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate: 0 (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deep: 6 (86)</td>
<td></td>
<td>7 (1/1)</td>
<td></td>
<td>X (7/7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F/U: 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pre indicates prepubertal girls (nonestrogenized hymen); Pub, pubertal girls (estrogenized hymen); F/U, follow-up examinations; —, no findings in this category; X, never seen.

^a Lacerations became shallower as they healed.
DISCUSSION

This study reconfirmed that female genital injuries heal rapidly and usually leave little if any evidence of the previous trauma. There was no significant difference in the rate of healing or the final outcome of the nonhymenal genital injuries in these 2 groups of girls (Tables 1–4). The search for a “pattern” or a “time sequence” in the healing process that could determine the age of a wound did not materialize. The amount of time required for an injury to resolve depended on its severity. In both groups, signs of the less severe injuries disappeared in 2 to 3 days, whereas evidence of the more severe injuries took up to 2 weeks to resolve. New blood vessel formation and the appearance of scar tissue were rare in both groups of girls; however, 2 findings that provided the examiner with a method for approximating the age of an injury were identified. One was the presence of petechiae, and the other was the detection of a blood blister. The presence of petechiae was an indication that the injury had occurred within the past 24 hours. None of these pinhead-sized lesions was found on the nonhymenal genital tissue beyond 24 hours in any of these girls (Figs 1–3). Petechiae were also detected on the hymen in this study’s companion report. Both studies used the same population of patients. In the companion report, petechiae were present on the hymen up to 48 hours in the prepubertal girls and up to 72 hours in the pubertal girls.

In both reports, there was a finding that initially seemed to be an exception to the time frame for the disappearance of a petechia. What originally were thought to be petechiae turned out to be small vascular anomalies. These lesions, which blanched with pressure, persisted for weeks.

The other finding was the blood blister (Fig 4). These blood-filled vesicles, which usually appeared for the first time during a follow-up examination, were associated with the more severe injuries. Their presence was an indication that an injury had occurred sometime in the past month. This finding was particularly helpful when all other signs of a recent injury had resolved.

The remarkable healing of the majority of these injuries was not unexpected. The healing process, as described by Kissane and Edwards and Dunphy, was similar to that found in this study. The less severe injuries, such as an abrasion or a superficial laceration, seemed to heal by regeneration in which all signs of the injury rapidly disappeared. The deeper lacerations most likely healed by repair. Only occasionally did any of these injuries leave behind new blood vessels (neovascular changes) or scar tissue formation (Table 4).

As a convenience-sample study, there are a number of limitations in this report. We did not have a predetermined protocol for the timing of the follow-up evaluations or specify the type of recording device to be used. Even the examination method was not mandated. In addition, not every injury was followed until complete healing had taken place. Although all of these

<table>
<thead>
<tr>
<th>Location</th>
<th>Neovascular, n/N (%)</th>
<th>Scar, n/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepubertal</td>
<td>0/8 (0)</td>
<td>0/8 (0)</td>
</tr>
<tr>
<td>Pubertal</td>
<td>0/8 (0)</td>
<td>0/8 (0)</td>
</tr>
<tr>
<td>Vestibule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepubertal</td>
<td>2/6 (33)</td>
<td>0/6 (0)</td>
</tr>
<tr>
<td>Pubertal</td>
<td>0/2 (0)</td>
<td>0/2 (0)</td>
</tr>
<tr>
<td>Fossa navicularis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepubertal</td>
<td>2/10 (20)</td>
<td>2/10 (20)</td>
</tr>
<tr>
<td>Pubertal</td>
<td>0/7 (0)</td>
<td>2/7 (29)</td>
</tr>
<tr>
<td>Posterior fourchette</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepubertal</td>
<td>1/14 (7)</td>
<td>7/14 (50)</td>
</tr>
<tr>
<td>Pubertal</td>
<td>0/4 (0)</td>
<td>1/4 (25)</td>
</tr>
<tr>
<td>Perineum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepubertal</td>
<td>0/8 (0)</td>
<td>6/8 (75)</td>
</tr>
<tr>
<td>Pubertal</td>
<td>0/2 (0)</td>
<td>0/2 (0)</td>
</tr>
</tbody>
</table>

New blood vessel and scar tissue formation were recorded only for patients who had been followed for ≥30 days.

FIGURE 4
Case 2: a 12-year-old who was sexually assaulted by 2 teenaged boys 24 hours before the current examination. Two blood blisters (arrows) in the midline of the vestibule can be seen (supine labial traction method).
factors were considered, we believed that to obtain a large enough sample size in a reasonable amount of time, we would need to obtain help from our colleagues throughout the United States. We decided to ask them to submit cases that they had previously evaluated.

The large number of cases submitted did provide us with an opportunity to evaluate both the healing processes and the outcome of most of these injuries. Although not all of the injuries had completely healed by the time of the last reevaluation, the majority of the findings had either disappeared or had significantly changed when last examined. The photographs obtained by the various recording devices proved to be excellent. In addition, the fact that half of the girls were examined by 3 different methods proved to be an unexpected and valuable finding. This discovery provided us with an opportunity to compare the results obtained through the use of the supine labial separation method, the supine labial traction technique, and the prone knee-chest position. The results of this comparison are reported in a companion article.25

The results from this project should help medical examiners interpret the anatomic findings of a girl's nonacute genital injury. The presence of petechiae or a blood blister should provide valuable information to the investigator who is attempting to determine when an injury might have occurred. The rapid and complete healing of the majority of these injuries should also prove to be reassuring to both the girl and her family.

CONCLUSIONS
The results from this study corroborate previous observations that nonhymenal genital injuries in a young girl or an adolescent heal remarkably well and tend to leave little, if any, evidence of the previous trauma. The rapid resolution of the petechiae along with the persistence of blood blisters did provide findings that can be used for approximating the age of an injury. Similar to the other types of injuries, the time required for resolution and the final appearance of the nonhymenal lacerations did depend on their severity. The rapid resolution of these injuries makes it imperative that a child or an adolescent be examined as soon as possible when there is a suspicion of a possible sexual assault.

APPENDIX: GLOSSARY

ACKNOWLEDGMENTS
This study was supported in part by the Lucile Packard Foundation; the Kenneth Jonsson Family Foundation; and the Children's Miracle Network through the Department of Pediatrics, University of California, Davis.

We gratefully acknowledge the participation of the following individuals in this multicenter, collaborative research project: David Kerns, MD (Valley Medical Center, San Jose, CA); Marilyn Kaufhold, MD (San Diego Children's Hospital); Mary Ritter, CHA (Valley Medical Center); Joan Voris, MD (University Hospital and Medical Center, Fresno, CA); Terrence Donald, MD (Women's and Children’s Hospital, North Adelaide, Australia); Martin Finkel, DO (University of Medicine and Dentistry of New Jersey, Stratford, NJ); Lori Frasier, MD (University Physicians-Green Meadows, Columbia, MO); Wendy Gladstone, MD (Exeter Pediatric Associates, Exeter, NH); Penny Grant (Child Abuse Network, Tulsa, OK); Nancy Harper, MD (Naval Medical Center, Portsmouth, VA); Roberta Hibbard, MD (Indiana University School of Medicine, Indianapolis, IN); Ralph Hicks, MD (Wright State University, Dayton, OH); Margie Hogan, MD (Hennepin County Medical Center, Minneapolis, MN); Michael Jordan, MD (Medical Office Building, Newark, NY); Michael Knappman, PA-C (Redwood Children's Center, Santa Rosa, CA); Kathie Kraley, RN, MSN, CPNP (Children's Mercy Hospital, Kansas City, MO); Susan Murawski, MS, ARNP, CPNP (Southwest Florida Children's Fund, Fort Myers, FL); Vincent Palusci, MD (Michigan State University College of Human Medicine, Grand Rapids, MI); Andrew SIrotnak, MD (Children's Hospital, Denver, CO); Naomi Sugar, MD (HCSATS, Seattle, WA); Andi Taroli, MD (Children’s Advocacy Center, Scranton, PA); Danny Waldrop, MD (Geisinger Medical Center, Danville, PA); and J. M. Whitworth, MD (Florida Children’s Fund, Jacksonville, FL).

Special thanks go to Ana Ross, PA-C, for assistance with reviewing the medical charts of the University of California Davis Medical Center patients; to Stephen Boos, MD, formerly a Lieutenant Colonel, United States Air Force Medical Corps, for review of the manuscript; and to Lisa Stenhouse and Lieschen Trelford for assistance with the myriad of tasks connected to the preparation of this manuscript.
REFERENCES

Twelve-Month Neurofunctional Assessment and Cognitive Performance at 36 Months of Age in Extremely Low Birth Weight Infants

Maria Lorella Gianni, MD, Odoardo Picciolini, MD, Chiara Vegni, MD, Laura Gardon, PT, Monica Fumagalli, MD, Fabio Mosca, MD

Neonatal Intensive Care Unit, Institute of Pediatrics and Neonatology, Fondazione IRCCS “Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena” University Medical School, Milan, Italy

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The objective of this study was to investigate whether an early neurofunctional assessment (at 12 months’ corrected age) is predictive of cognitive outcome at 36 months of age in extremely low birth weight infants.

METHODS. We conducted an observational longitudinal study. Neurodevelopmental outcome by means of a neurofunctional assessment was evaluated at 12 months’ corrected age and 36 months’ chronological age in 141 extremely low birth weight children. Cognitive outcome was assessed with use of the Griffiths Mental Developmental Scale.

RESULTS. A significant association was found between the 12-month neurofunctional status and cognitive performance at 36 months. A higher general quotient on the Griffiths Mental Developmental Scale at 36 months was observed in infants who exhibited normal (score: ≥1) neurodevelopment compared with children who exhibited minor (score: 2) and major (score: ≥3) dysfunctions at the 12-month neurofunctional evaluation (99 ± 6.8 vs 85.3 ± 16.3 vs 57.3 ± 22.0). A score of ≥2 at the 12-month neurofunctional assessment, abnormal brain MRI results at term, and chronic lung disease remained predictive of cognitive delay at 36 months of age and also after adjustment for confounders.

CONCLUSIONS. The 12-month neurofunctional evaluation may be an additional useful clinical tool in predicting later cognitive outcome in extremely low birth weight children.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-3364
doi:10.1542/peds.2006-3364

Key Words
neurofunctional assessment, cognitive outcome, extremely low birth weight infants

Abbreviations
VLBW—very low birth weight
ELBW—extremely low birth weight
SGA—small for gestational age
NEC—necrotizing enterocolitis
ROP—retinopathy of prematurity
CLD—chronic lung disease
HC—head circumference
GQ—general quotient

Accepted for publication May 29, 2007

Address correspondence to Maria Lorella Gianni, MD, Neonatal Intensive Care Unit, Clinica Mangiagalli, IRCs, University Medical School, Via Commenda 12, 20122 Milano, Italy. E-mail: maria.gianni@unimi.it

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
Survival of very low birth weight (VLBW; <1500 g) infants has recently improved as a result of advances in perinatal care. Major disabilities occur in 15% to 20% of the VLBW population, with a higher prevalence in infants with the youngest gestational age and the lowest birth weight. A high rate of learning difficulties, behavioral problems, and lower IQ scores, compared with the population mean, is reported in survivors without major dysfunctions, especially extremely low birth weight (ELBW; <1000 g) infants. Early detection of infants who are at high risk for poor neurodevelopmental outcome remains a challenge for clinicians to make referral for intervention and optimize potential outcome. Several authors have investigated the relationship between motor ability, mainly investigated by a developmental assessment of movement, and cognitive performance in ELBW, suggesting that early movement problems may be predictive of later cognitive adverse outcome even in the absence of defined neurologic problems.

We previously reported that an early neurofunctional evaluation in VLBW infants may be an additional useful tool in reassuring parents on the integrity of central nervous system function. Few data are available in the literature on any hypothetical relationship between cognitive outcome of ELBW infants and early neurofunctional assessment. The aim of this study was to examine whether a neurofunctional evaluation at 12 months is predictive of cognitive outcome at 36 months of age in ELBW infants.

METHODS
Among all consecutive newborn infants who were admitted at the same Institution from 1996 to 2001, 159 infants entered the study. Inclusion criterion was birth weight ≤1000 g. Exclusion criteria were presence of congenital diseases and/or chromosomal abnormalities or infant death during postpartum hospital stay. Infants were scheduled to be prospectively followed up to 36 months of age. The Consolidated Standards of Reporting Trials (CONSORT) flowchart of the study is shown in Fig 1. Written informed consent was obtained from the newborns’ parents, and the departmental ethics committee approved the study design.

Presence of gestational hypertension and/or pre-eclampsia, steroids received before delivery, and educational level were investigated as maternal variables. Gestational hypertension and preeclampsia were defined, respectively, as de novo hypertension (systolic blood pressure ≥ 140 mm Hg or diastolic BP ≥ 90 mm Hg) arising after midpregnancy and gestational hypertension associated with new-onset proteinuria (≥300 mg/24 hours). The educational level of mothers was evaluated and categorized as low (≤13 years) or high (>13 years).

Neonatal characteristics (gestational age, being appropriate or small for gestational age [SGA], birth weight, length, and head circumference; need for mechanical ventilation; and the occurrence of sepsis, necrotizing enterocolitis (NEC; classified as stage 3 according to the classification of Bell et al), bronchopulmonary dysplasia, and retinopathy of prematurity (ROP) of stage 3 or higher were recorded prospectively. Gestational age was based on the last menstrual period and first-trimester ultrasonogram. Infants with birth weights of ≥10th or <10th percentile for gestational age, according to the North-Italian growth charts, were classified as appropriate for gestational age or SGA, respectively. Sepsis was defined by the presence of a positive blood culture. Patients with NEC were pooled together with those who

![CONSORT flowchart](https://example.com/fig1.png)

**FIGURE 1**
Consolidated Standards of Reporting Trials (CONSORT) flowchart.
had sepsis because of the strong association of NEC with infection. Chronic lung disease (CLD) was defined by use of supplemental oxygen at 36 weeks’ postconceptional age. Corrected age was calculated, up to 24 months of life, from the chronologic age adjusting for gestational age. Brain MRI was performed at a mean of 40 ± 2 weeks’ postconceptional age in all newborns. An abnormal MRI result was defined by the presence of major brain lesions (eg, ventriculomegaly, cystic and noncystic periventricular leukomalacia, focal parenchymal brain lesions).

Infants entered a follow-up program that consisted of periodic pediatric visits at 3, 6, 12, and 24 months’ corrected age and 36 months’ chronological age and of evaluation of neurodevelopmental outcome by means of the neurofunctional evaluation at 12 months’ corrected age and 36 months’ chronological age. Cognitive outcome was assessed with use of the Griffiths Mental Developmental Scale at 36 months’ chronological age.

Weight, length, and head circumference (HC) were measured using standardized procedures. Growth z scores were then calculated by EuroGrowth 2000 software (EuroGrowth Study Group, Vienna, Austria). Catch-up growth was defined by a growth z score of more than −2 SD.

The neurofunctional evaluation was based on the study of evoked and spontaneous motility, postural adaptability, variability of motor patterns, and neuromotor and behavioral skills. A detailed description of the neurofunctional assessment is reported in the Appendix. Infants were categorized into 3 groups, according to the neurofunctional status, as normal (score: 0–1), exhibiting minor dysfunctions (score: 2), or exhibiting major dysfunctions (score: 3–4). The Griffiths Mental Developmental Scale and related subscales (locomotor, personal and behavioral skills) were administered by the same expert trained examiner, who was blind to the neurofunctional evaluation. A general quotient (GQ) was then calculated. The mean value of GQ was 100 with an SD of 12. A GQ of <70 was classified as severe developmental delay.

**Statistical Analysis**

Descriptive data are shown as means ± SD or number of observations (percentage). Comparison among groups was performed by the χ² test for discrete variables or by analysis of variance for continuous variables. Significance of multiple comparisons was adjusted by the least significant difference test correction. Differences within patients in repeated measurements of growth parameters were assessed by analysis of variance. Logistic regression analysis was used to identify determinants of developmental delay (GQ < 88) at 36 months of age. Maternal and neonatal characteristics that are known to be associated with adverse cognitive outcome were entered in the model. Factors examined included maternal education, prenatal steroids, maternal hypertension, multiple birth, gender, birth weight, gestational age, being SGA, sepsis, ROP of stage 3 or higher, HC catch-up growth at 12 months, neurofunctional status at 12 months’ corrected age, CLD, and cranial MRI status at term.

All statistical analyses were conducted at the α = .05 level and were 2-tailed. Statistical analysis was performed by using SPSS 12 (SPSS Inc, Chicago, IL).

**RESULTS**

Follow-up data at 36 months’ chronological age were available for 141 (70 girls; 71 boys) infants. One infant died through the follow-up. No differences in the characteristics at birth and in the developmental measures, when last assessed, were observed between infants who were lost at follow-up and those who were evaluated. The characteristics of the studied population at birth are shown in Table 1.

Maternal hypertension occurred in 35.1% and 39.4% of mothers with infants who exhibited a GQ of <88 and ≥88, respectively, at 36 months’ chronological age (P = .39). Antenatal steroids were administered to 50% and 59.5% of mothers with infants who exhibited a GQ of <88 and ≥88, respectively, at 36 months’ chronological age (P = .2). Maternal educational level was low in 73.1% and 83.8% of mothers with infants who exhibited a GQ of <88 and ≥88, respectively, at 36 months’ chronological age (P = .2). The characteristics of the studied population according to cognitive outcome at 36 months’ age are shown in Table 2.

Younger gestational age, being male, ROP of stage 3 or higher, and CLD were associated with a GQ of <88 at 36 months of age. A GQ of <70 was found in 14.2% (n = 20) of children. Abnormal MRI status was found in 72.4% and in 27.6% of infants who exhibited a GQ of <88 and ≥88, respectively, at 36 months’ chronological age (P < .0001).

Mean (SD) weight, length, and HC z scores ranged through the first 36 months of life from −2.62 (1.7) at 3 months to −1.77 (1.3) at 36 months (P < .0001), from −2.67 (1.5) to −1.01 (1.07; P < .0001), and from −1.57 (1.38) to −1.57 (1.03; P = 0.9), respectively. At 36 months, 38%, 17%, and 34% of infants had weight,
length, and HC z score less than −2 SD, respectively. Absence of HC catch-up growth at 36 months was associated with a GQ of <88 at 36 months (P < .05).

Neurofunctional Assessment at 12 Months’ Corrected Age and at 36 Months’ Chronological Age

At 12 months’ corrected age, the neurofunctional score was ≤1 in 82 (58.2%) infants, 2 in 41 (29%), and ≥3 in 18 (12.8%). The corresponding values at 36 months’ chronological age were 67 (47.5%), 52 (36.9%), and 22 (15.6%) infants. At the age of 36 months, the neurofunctional status improved, unchanged, or deteriorated in 7 (5%), 108 (76.6%), and 26 (18.4%) infants, respectively; 73% and 100% of infants with a score of ≤1 or ≥3, respectively, at 12 months’ corrected age had unchanged functional status at 36 months of age. The clinical disabilities of infants who had a neurofunctional score of 2 at 36 months of age were learning disabilities and dyspraxia (23%), clumsiness (23%), behavioral problems (30%), language impairment (10%), psychomotor retardation (12%), and hypoacusia (2%). The clinical features of the infants who had a neurofunctional score of ≥3 at 36 months of age were cerebral palsy (63%), mental retardation (22%), visual impairment (5%), deafness (5%), and autism spectrum (5%).

Neurofunctional Status at 12 Months’ Corrected Age and Griffiths Mental Development Scale at 36 Months’ Chronological Age

A significant statistical association between the 12-month neurofunctional status and the GQ assessed at 36 months (P < .0001) was found. On posthoc analysis, all comparisons were significant.

Predicting GQ at 36 Months’ Chronological Age From the 12-Month Neurofunctional Assessment

A logistic regression analysis was performed to identify determinants of developmental delay at 36 months (Table 3). For statistical analysis, patients who had a score of ≥2 at the 12-month neurofunctional evaluation were pooled. Abnormal cranial MRI results at term, CLD, and a score of ≤1 at the neurofunctional evaluation at 12 months were found to be independently associated with cognitive delay at 36 months.

DISCUSSION

In this study, the 12-month neurofunctional assessment, MRI status at term, and CLD remained predictive of cognitive development at 36 months in ELBW infants, also after adjusting for confounders. In addition, we found that children who were assessed as normal (score ≤ 1) at the 12-month neurofunctional assessment showed significantly higher GQ at 36 months of age compared with infants who exhibited dysfunctions (score ≥ 2). This finding could be explained by the fact that the detection of dysfunctions at 12 months of age may reflect the difficulty that the infant experiences in the learning process and in the development of new tasks, as a result of impairment of the emerging functions at that age, evaluated by the neurofunctional approach. As a consequence, the normal process of cognition may be impaired.

Several authors have found an association between motor function assessed at 12 months and cognitive outcome at school age. Burns et al4 reported that group classification of motor development at age 1 year, investigated using the neurosensory motor developmental assessment, is predictive of cognitive outcome in ELBW infants at age 4 years. Jeyaseelan et al3 reported that motor difficulties, assessed by use of neurosensory motor developmental assessment, in ELBW infants at 2 years are strongly associated with later clinical measures of attention. Roth et al15 and Stewart et al14 found a close relationship between neuromotor impairment, investigated by standardized neurologic examination, and cognitive defects at age 4 years. The results of these studies are consistent with Piaget’s theory15 and the mirror neuron system.16,17 According to Piaget, there is a strict interrelation between motor and cognitive development: action is considered the foundation of knowledge as it is,

### TABLE 2  Characteristics of the Studied Population According to Cognitive Outcome at 36 Months of Age

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cognitive Assessment at 36 mo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GQ &lt; 88 (n = 37)</td>
</tr>
<tr>
<td>Birth weight, mean (SD), g</td>
<td>827.8 (143.0)</td>
</tr>
<tr>
<td>Birth length, mean (SD), cm</td>
<td>33.6 (2.6)</td>
</tr>
<tr>
<td>Birth HC, mean (SD), cm</td>
<td>24.1 (1.9)</td>
</tr>
<tr>
<td>Gestational age, mean (SD), wk</td>
<td>27.2 (1.8)</td>
</tr>
<tr>
<td>Males, n (%)</td>
<td>24.0 (64.9)</td>
</tr>
<tr>
<td>SGA, n (%)</td>
<td>10 (27.0)</td>
</tr>
<tr>
<td>Multiple birth, n (%)</td>
<td>12 (32.4)</td>
</tr>
<tr>
<td>Sepsis, n (%)</td>
<td>8 (22.9)</td>
</tr>
<tr>
<td>ROP stage 3 or higher, n (%)</td>
<td>10 (27.8)</td>
</tr>
<tr>
<td>CLD, n (%)</td>
<td>25 (67.6)</td>
</tr>
</tbody>
</table>

NS indicates not significant.

* P or χ² test.
TABLE 3  Variables Associated With a GQ of <88 at 36 Months of Age: Logistic Regression Analysis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Adjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestational age (≤28 vs ≥28 wk)</td>
<td>1.70 (0.30–8.40)</td>
</tr>
<tr>
<td>Birth weight (&lt;800 vs ≥800 g)</td>
<td>2.70 (0.50–14.00)</td>
</tr>
<tr>
<td>Gender (male vs female)</td>
<td>1.40 (0.40–5.40)</td>
</tr>
<tr>
<td>Being SGA (yes vs no)</td>
<td>0.70 (0.10–4.70)</td>
</tr>
<tr>
<td>Multiple birth (yes vs no)</td>
<td>0.90 (0.20–4.60)</td>
</tr>
<tr>
<td>Sepsis (yes vs no)</td>
<td>0.50 (0.10–2.40)</td>
</tr>
<tr>
<td>ROP stage 3 or higher (yes vs no)</td>
<td>0.60 (0.08–3.90)</td>
</tr>
<tr>
<td>CLD (yes vs no)</td>
<td>8.80 (1.80–42.90)</td>
</tr>
<tr>
<td>HC z score at 12 mo (absence vs presence of catch-up growth)</td>
<td>2.30 (0.60–9.70)</td>
</tr>
<tr>
<td>Neurofunctional status at 12 mo (normal vs impaired)</td>
<td>0.15 (0.03–0.70)</td>
</tr>
<tr>
<td>MRI status at term (abnormal vs normal)</td>
<td>16.70 (3.30–84.30)</td>
</tr>
<tr>
<td>Maternal hypertension (yes vs no)</td>
<td>0.80 (0.20–3.70)</td>
</tr>
<tr>
<td>Prenatal steroids (yes vs no)</td>
<td>0.70 (0.20–2.90)</td>
</tr>
<tr>
<td>Maternal education (low vs high)</td>
<td>1.60 (0.30–8.40)</td>
</tr>
</tbody>
</table>

OR indicates odds ratio; CI, confidence interval.

a P < .05.
b P < .005.

at first, physical and then it is transformed into mental action and representation. Also, the mirror neuron system, which activates when an individual performs an action, as well as when he or she observes a similar action done by another individual, seems to play an important role in the social cognition process: it seems not only to provide an action recognition mechanism but also to constitute a neural system for coding the intentions of others and allowing imitation learning. In addition, according to Hadders-Algra, both primary (the phase in which the variation in motor behavior is not geared to external conditions) and secondary (the phase in which the motor performance can be adapted to specific situations) variability and, consequently, the motor repertoire can be reduced in preterm infants as a result of perinatal lesions or abnormal cerebral cortical development. Consistent with these findings, the association of the 12-month neurofunctional evaluation with cognitive performance at 36 months of age may be to some extent explained by the fact that a neurofunctional approach includes the evaluation of evoked and spontaneous motility, postural adaptability, and variability of motor patterns, according to the emerging functions of each age. The finding of lower GQ at 36 months in infants who exhibited minor dysfunctions at the 12-month neurofunctional assessment highlights the importance of additional implementation of individualized care plans both during hospital stay and in the postdischarge period because a limited motor repertoire, even in the absence of brain lesions, may result in deficit of processing information and solving problems through the selection of the most appropriate strategies and adaptive solutions. In addition, promoting the infant’s neurobehavioral together with motor organization has been reported to be associated with both medical and developmental advantages; therefore, detection of dysfunctions at the 12-month neurofunctional evaluation may represent an additional clinical tool in identifying infants who are at risk for poor later cognitive outcome and could benefit from early intervention.

Although efforts have been made to eliminate the artificial division between the neurologic and the behavioral approach to the neurodevelopmental assessment of preterm infants, few studies have investigated the relationship between the cognitive outcome of ELBW infants and the neurofunctional evaluation. Wallace et al reported that preterm neonates who had poor neurobehavioral performance, assessed as visual-following and auditory-orienting, had significantly lower cognitive test scores at age 1 and 6 years. Costandinou et al found that the Neurobehavioral Assessment of the Preterm Infants, evaluated before hospital discharge, correlates with cognitive outcome at 30 months of age.

There is increasing evidence that cognitive domain in preterm infants is one of the most common areas of poor functioning. The prevalence of severe developmental delay found in our study is 14.2%. Several multicenter trials reported different rates of developmental disabilities in very preterm infants, ranging from 15% to 20% up to 35% to 40%. In a meta-analysis, Bhutta et al reported that mean cognitive scores of preterm infants at school age were 10.9 lower compared with those of control subjects, showing a linear correlation with birth weight and gestational age. In this study, a younger gestational age, being male, absence of HC catch-up growth at 36 months, and severe ROP were associated with developmental delay at 36 months at univariate analysis. No association with birth weight was found, probably because in our study, infants with birth weight <600 g were unrepresented. Being male, having poor postnatal growth, especially of the head, and having severe ROP have been reported as risk factors for developmental delay by other authors.

In this study, at logistic regression analysis, also abnormal MRI status at term and the presence of CLD remained predictive of cognitive outcome at 36 months. Abnormal MRI status at term has been reported as a strong predictor of adverse neurodevelopmental outcome by other authors, suggesting a role for MRI at term in risk stratification for preterm infants. There is agreement on the unfavorable effect of CLD, even in the absence of severe brain lesions, on cognitive outcome in children who are born very preterm, which seems to persist into school age.

In our study, maternal education and being SGA were not predictive of later cognitive outcome. The effect of maternal education on cognitive outcome in ELBW infants remains in question. Although several authors have reported lack of maternal education as a risk factor for poor cognitive development, others did not observe any effect. Current literature is inconclusive also about
the role of being SGA in determining neurodevelopment outcome. Some studies\(^1\) reported no effect of SGA status on cognitive outcome, whereas others found a modest association between SGA and IQ.\(^2\)

We did not find any association between sepsis and developmental delay. This finding is in agreement with Vohr et al.,\(^2\) who did not identify sepsis as predictive of cognitive impairment, whereas Stoll et al.\(^3\) found a rate of cognitive developmental disability of 33% to 42% in infants with infection versus 22% in infants without infection.

A potential limitation of the findings of this study is the loss to follow-up, which was >10%; however, because the patients who were lost to follow-up had similar characteristics at birth and in the developmental measures, when last assessed, to those reported, this loss is unlikely to bias substantially the associations reported. An additional limitation of this study was that, despite a relatively large sample size, the low rate of severe cognitive impairment precluded a separate analysis to identify the variables that were predictive of a GQ of <70.

Because disability, defined as impairment of function, has been recognized to be multidimensional,\(^4,5\) the choice of outcome classifications and assessments remains difficult. Additional investigations are warranted to identify additional, simple methods of neurodevelopmental assessment in clinical practice that may better reflect the emerging concept of disability and focus the attention on the consequences of disease rather than on disease itself. A functional approach to the neurodevelopmental evaluation that actually investigates not only an infant’s motor but also neuromotor and behavioral skills may well relate to later outcomes and focus on the outcome potential of the infant.

**CONCLUSIONS**

According to our findings, the presence of a score of ≥2 at the 12-month neurofunctional assessment represents an early clinical marker of adverse later cognitive outcome. This study suggests that an early neurofunctional evaluation could be an additional useful tool in alerting clinicians to follow up neurodevelopment strictly to initiate early intervention programs. Additional, larger studies are desirable to clarify better the role of the neurofunctional assessment in the early identification of children who are at risk for poor later cognitive performance.

**REFERENCES**


## APPENDIX: NEUROFUNCTIONAL ASSESSMENT OF PREMATURE INFANT AT 12 MONTHS

### Classification at 12 months

0. **typical result:** complete patterns of movement and interaction

1. slight anomalies which normalize with facilitation during the exam

2. evident abnormal result: the anomalies remain during examination: the function (motor, postural, adaptive) is moderately troubled, but possible (not prevented)

3. the function is difficult: the anomalies significantly disturb the function

4. severe anomalies that upset or prevent the function, till fixed pathological patterns of movement

### Time:
20 minutes

### POSTURAL ANOMALIES

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Slight</td>
</tr>
<tr>
<td>1</td>
<td>Marked</td>
</tr>
<tr>
<td>2</td>
<td>None</td>
</tr>
</tbody>
</table>

### ITEMS

<table>
<thead>
<tr>
<th>ITEMS</th>
<th>DESCRIPTION</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ADAPTENESS &gt; AUTONOMY</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attention-interaction</td>
<td>stability</td>
<td></td>
</tr>
<tr>
<td>Feeding function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastrointestinal function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep-wake rhythms</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NEUROSENSORY FUNCTION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hearing and language</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BEHAVIOURAL FUNCTION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social interaction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional stability (or</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hyperactivity)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal social behaviour</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SPONTANEOUS MOTOR ACTIVITY</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor skills</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>POSTURAL PATTERNS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sitting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standing</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TEMPORAL PATTERNS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balancing (sitting)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squatting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equilibration (kneeling,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>standing)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral abduction reaction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>of hip</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SEQUENCES</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axial rotation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antigravity rotation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rolling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creeping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rightings</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FACILITATION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower limb abduction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hand opening</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PRAXIES (FINE- GRAPHO- MOTOR)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tower building</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shapes matching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Action patterns”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of play</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ANGULAR MEASURES</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adductors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ischiobiales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sural triceps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper limb extension</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PRIMITIVES&gt;PATHOLOGICAL REFLEXES</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymmetrical tonic neck</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plantar grasp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Babinski (evoked or</td>
<td></td>
<td></td>
</tr>
<tr>
<td>spontaneous)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Shorter Sleep Duration Is Associated With Increased Risk for Being Overweight at Ages 9 to 12 Years

Julie C. Lumeng, MDa,b, Deepak Somashekar, BSa, Danielle Appugliese, MPHc, Niko Kaciroti, PhD,a, Robert F. Corwyn, PhDd, Robert H. Bradley, PhD*e

aCenter for Human Growth and Development and bDepartment of Pediatrics and Communicable Diseases, University of Michigan, Ann Arbor, Michigan; cData Coordinating Center, Boston University, Boston, Massachusetts; dDepartment of Psychology and eCenter for Applied Studies in Education, University of Arkansas, Little Rock, Arkansas

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The potential association between short sleep duration or sleep problems and childhood overweight has not been well described. The objective of this study was to test the independent associations of sleep duration and problems with overweight risk in children.

METHODS. Data from the National Institute of Child Health and Human Development Study of Early Child Care and Youth Development were analyzed. In 3rd and 6th grades, sleep duration and problems were obtained by maternal report, and height and weight were measured, with overweight defined as a BMI of ≥95th percentile for age and gender. Logistic regression evaluated the association of sleep duration and problems with overweight at 6th grade cross-sectionally adjusting for gender, race, and maternal education. Additional covariates tested individually included the level of chaos at home, the quality of the home environment, the lax-parenting subscale score of the Raising Children Checklist, and the Child Behavior Checklist internalizing and externalizing subscale scores. Logistic regression also evaluated the relationship of sleep duration at 3rd grade and overweight at 6th grade, adjusting for gender, race, maternal education, and the child’s BMI z score in 3rd grade.

RESULTS. Of 785 children, 50% were male, 81% were white, and 18% were overweight in 6th grade. Shorter sleep duration in 6th grade was independently associated with a greater likelihood of overweight in 6th grade. Shorter sleep duration in 3rd grade was also independently associated with overweight in 6th grade, independent of the child’s weight status in 3rd grade. Sleep problems were not associated with overweight.

CONCLUSION. One preventive approach to overweight may be to ensure adequate sleep in childhood.
Emerging research has revealed sleep to be an important regulator of many physiologic functions, including energy balance, appetite, and weight maintenance.\(^1\) The relationship between sleep and weight has become a topic of great interest as US obesity rates reach record levels\(^3\) and chronic sleep deprivation affects a growing percentage of American youth and adults.\(^4\) Several studies have demonstrated that even modest reductions in sleep duration are associated with significant increases in obesity risk among adults.\(^1-4\) Sleep curtailment has also been linked to alterations in leptin and ghrelin levels and impaired glucose tolerance, suggesting that long-term reductions in sleep may set up hormonal changes that lead to weight gain.\(^2,9\)

The link between sleep duration and obesity has been well established in adults,\(^5-7\) but comparatively little is known about this relationship in adolescents and younger children. Several studies\(^10-12\) involving 8000 Japanese children of the Toyama Birth Cohort showed an association between shorter sleep duration and increased overweight risk in age ranges of 3 to 7 years. Similar findings were reported using the Avon Longitudinal Study of Parents and Children,\(^13\) a cohort of British children (96% white) that examined sleep duration at 30 months and overweight risk at 7 years. Three studies\(^14-16\) have demonstrated this relationship in US children. Two of these demonstrated a concurrent association in racially diverse cohorts, although neither evaluated the potential confounding role of socioeconomic status (SES). Among 60 overweight children aged 10 to 17 years compared with matched control subjects, overweight was associated with significantly shorter sleep duration,\(^14\) and among 383 children aged 11 to 16 years, decreased sleep duration was associated with increased overweight risk, independent of age, gender, and race.\(^15\) The single US study that evaluated the potential longitudinal relationship between sleep and overweight found that among 150 children (primarily white and with well-educated parents), shorter sleep duration between the ages of 2 and 5 years was associated with increased overweight risk at age 9.5 years.\(^16\)

Collectively, these studies provide evidence for an association between shorter sleep duration and an increased likelihood of overweight in children, but their interpretation is limited by racial and/or socioeconomic homogeneity within the US cohorts or a lack of longitudinal data. Children’s reported sleep habits and sleep problems differ by race and SES.\(^17-20\) Black children are reported to nap more frequently and until older ages than white children but get less sleep at night.\(^18\) Hispanic adolescents are reported to have more insomnia than other ethnic groups,\(^19\) and children of lower SES have later sleep onset and shorter sleep duration than children of higher SES.\(^17,21\) Because the prevalence of overweight in the United States also differs significantly by race and SES,\(^22-24\) the association between sleep duration and overweight risk may be confounded by these characteristics. Our primary aim was to test the hypothesis that shorter sleep duration is associated with increased likelihood of overweight in US grade-school children independent of race and SES.

A secondary goal of this study was to evaluate the potential association between reported sleep problems and childhood overweight risk. Poor sleep quality has been associated with a lowered sense of well-being and decreased quality of life in young adults,\(^25\) but its role in obesity is not well understood. Several studies\(^26,27\) have shown a link between obesity and obstructive sleep apnea in both adults and children; however, the highly specific nature of sleep apnea makes it difficult to extrapolate such findings to the report of more general sleep problems such as night waking, delayed sleep onset, and restlessness. Two studies with sample sizes of fewer than 500 subjects evaluated general sleep quality by using questionnaires in adults\(^7\) or wrist actigraph measurements in children/adolescents\(^15\) and found no association between poor sleep quality and obesity; however, a recent study\(^14\) that used both questionnaire and actigraphy in 60 children between the ages of 10 and 17 years found a positive association between sleep disturbance and overweight. More recent work indicated that excessive daytime sleepiness in adults is associated with a higher prevalence of obesity, independent of obstructive sleep apnea and sleep duration.\(^28\) We sought to test the hypothesis that the report of more sleep problems would be associated with increased risk for overweight in grade-school children.

Finally, we sought to determine whether the relationship between short sleep duration and increased likelihood of overweight, if identified, persisted when controlling for measures of quality of the home environment, parenting, and child behavior problems. We hypothesized that an identified relationship between shorter sleep duration and overweight may simply reflect an underlying lack of structure or suboptimal parenting in the home and therefore sought to test the potential confounding role of these factors. Sleep disturbance is a defining feature of affective disorders,\(^29\) and behavior problems and affective disorders have been associated with an increased risk for overweight in children.\(^30-32\) We hypothesized that behavior problems may underlie both short sleep duration and increased risk for overweight and therefore also sought to test this factor as a potential confounder.

**METHODS**

The sample was composed of children and their parents who were enrolled in the National Institute of Child Health and Human Development Study of Early Child Care and Youth Development (NICHD-SECCYD),\(^33\) a longitudinal study of relations between child behavior and development, particularly in relation to child care...
experience. Families were recruited shortly after the birth of a child in 1991 from 10 areas of the United States, both urban and rural, and data were collected prospectively from birth onward. Details of the recruitment method and sampling plan are available elsewhere. The initial sample included 1364 children and was representative of the demographics of the catchment area from which the sample was recruited. By 6th grade, 1077 children were still enrolled in the study, and 806 of them had anthropometric data in 6th grade. Much of the missing anthropometric data were accounted for by families moving to other communities, which prevented them from attending the laboratory visits where the measurements were taken but allowed continued participation in questionnaires and telephone interviews. Children with missing anthropometric or sleep-duration data at 6th grade were excluded from the analysis, which resulted in a final sample size of 785 children (58% of the original cohort). This study was approved by the institutional review boards of all participating institutions.

**Measurement of Sleep**

Data regarding sleep duration and sleep problems were obtained by maternal report on the Children’s Sleep Habits Questionnaire (CSHQ) in 3rd and 6th grades, when children were on average 9.02 (SD: 0.31) and 11.61 (SD: 0.15) years of age, respectively. The CSHQ is a validated parent questionnaire regarding the child’s sleep duration and problems, modified for use in the NICHD-SECCYD to consist of 28 items. Sleep duration was based on maternal response to the question, “How much sleep does your child get each day (including naps)?” Mothers reported sleep characteristics on the remaining 27 items (Appendix) using a 3-point scale: 1 = usually, 2 = sometimes, and 3 = rarely). Four summary scores were computed by the NICHD-SECCYD as the average of their contributing question items: bedtime problems, night waking problems, morning waking problems, and daytime sleepiness (Cronbach $\alpha = .60$, .64, .74, and .65, respectively). We created a general sleep problems score by taking the mean of the 4 subscores. The possible range was therefore 1 to 3, with higher scores indicating more sleep problems.

**Measurement of Overweight**

Height and weight were measured during laboratory visits in 3rd and 6th grades by trained research assistants using a protocol standardized at all 10 sites. BMI was calculated and child overweight defined dichotomously as a BMI of $\geq 95$th percentile for age and gender based on National Center for Health Statistics norms. BMI $z$ score at 3rd grade was also calculated using these norms.

**Covariates**

Gender was included as a covariate given the association of male gender with shorter sleep duration and increased overweight prevalence. Race (categorized as white versus nonwhite) was included as a covariate because of the association of minority race/ethnicity with increased childhood overweight prevalence and sleep problems. Race was categorized in this manner because of the small number of children of each race/ethnicity besides white in this cohort. SES, indexed as years of maternal education, was included because of the association of lower SES with increased overweight prevalence, shorter sleep duration, and more sleep problems. Maternal education is correlated with income-to-needs ratio (family income relative to the poverty line, accounting for family size) in this data set (Spearman rank correlation $= 0.59, P < .001$). Maternal education was used as the primary index of SES given that it provided a larger sample size.

For determination of whether an identified relationship between shorter sleep duration and overweight simply reflected an underlying lack of structure or suboptimal parenting, 3 additional covariates were considered. The CHAOS Scale is a 15-item validated questionnaire that is completed by mothers in 3rd grade to assess the degree of environmental chaos in the home. Items are responded to as “true” (1 point) or “false” (2 points), thus resulting in a possible range of 15 to 30, with higher scores indicating more chaos in the home. The quality of the home environment in 3rd grade was measured by the Mid-Childhood Home Observation for Measurement of the Environment (HOME), one of the most widely used indices of the quality and quantity of stimulation and support available to a child in the home. Information is obtained during a home visit via observation and interview. It is composed of 55 items, each of which is scored in a binary manner (yes/no), with scores therefore ranging from 0 to 55 and higher scores indicating higher quality home environments. The HOME has consistently been correlated with cognitive, language, achievement, and socioemotional outcomes. For assessment of parental discipline strategies, mothers completed a questionnaire that was adapted from Greenberger’s Raising Children Checklist when the child was in 3rd grade. We included the lax control subscale, which consists of the sum of responses to 9 items. Response choices ranged from 1 to 4 (“definitely no” to “definitely yes”), and subscale scores therefore ranged from 9 to 36, with higher values indicating more lax parenting (Cronbach $\alpha = .73$).

Child behavior problems were assessed in 3rd grade via the Child Behavior Checklist (CBCL), a 99-item rating scale that is the most widely used assessment of behavioral problems in children. Scores are presented as T scores, which have a mean of 50 and an SD of 10. A cutoff of 60 is frequently used to denote clinically signif-
significant behavioral problems. We specifically tested the internalizing and externalizing subscale scores as covariates, given that sleep disturbance is a defining feature of affective disorders and the reported association of behavior problems and affective disorders with overweight in children.

Evaluation of Missing Data
We compared the sample with complete data for overweight status and sleep duration at 6th grade (n = 785) with the sample without complete data for these 2 variables. Children without complete data had mothers with fewer years of education (mean [SD]: 14.0 [2.6] vs 14.4 [2.4]; P = .02). There was no difference in gender (P = .16) or race/ethnicity (P = .61).

Statistical Analyses
Analyses were conducted with SAS 9.1 (SAS Institute, Cary, NC). The primary outcome of interest was overweight status (yes versus no) in 6th grade. We first performed unadjusted bivariate analyses (t test for continuous variables and χ² for categorical variables) comparing each of our main effect sleep measures as well as each of our covariates (gender, race, maternal education, CHAOS score, HOME total score, lax-parenting subscale score, and CBCL internalizing and externalizing subscale scores) by the child’s overweight status in 6th grade. We also evaluated the relationship between each of these covariates and mother-reported sleep duration and problems at 6th grade using t tests and correlation coefficients.

Adjusted Concurrent Relationship: Sleep Duration and Overweight at 6th Grade
A multiple logistic regression was used to test the independent relationship of sleep duration with overweight status concurrently at 6th grade, controlling for gender, race, and maternal education. The model was repeated to test the following covariates individually in the model one by one: mother-reported bedtime at 6th grade, mother-reported wake time at 6th grade, mother-reported sleep problems at 6th grade, CHAOS score at 3rd grade, HOME total score at 3rd grade, lax-parenting subscale score at 3rd grade, and CBCL internalizing and externalizing subscale scores at 3rd grade.

Longitudinal Relationship: Sleep Duration Between 3rd and 6th Grades and Overweight at 6th Grade
To test the longitudinal relationship of sleep duration in 3rd and 6th grades with overweight status at 6th grade, we created a model that included the main predictors of sleep duration in 3rd grade and change in sleep duration between 3rd and 6th grades. Covariates included gender, race, and maternal education. The child’s BMI z score in 3rd grade was also included as a covariate to assess whether the relationship between sleep duration from 3rd to 6th grade and overweight at 6th grade was present independent of the child’s weight status at the beginning of this time frame. The presence of a relationship while controlling for the child’s weight status at the beginning of the period would provide evidence supporting a causal relationship between short sleep duration and increased likelihood of future overweight. We also tested the quadratic and cubic polynomial terms for the “change in sleep duration” variable and “change in sleep duration” in tertiles (tertile 1: a decline in sleep duration ≥1.6 hours [9.2%]; tertile 2: change in sleep duration between −1.5 and 0.3 hours [78.3%]; and tertile 3: an increase in sleep duration >0.3 hours [12.3%]).

Sleep Problems at 3rd and 6th Grades and Overweight at 6th Grade
Two logistic regression models were tested to evaluate the relationship between sleep problems and overweight at 6th grade. One model tested the general sleep problems score at 3rd grade as the main predictor of overweight at 6th grade controlling for gender, race, and maternal education. The second model tested the general sleep problems score at 6th grade as the main predictor of overweight at 6th grade while again controlling for gender, race, and maternal education. Adjusted odds ratios (ORs) and their corresponding 95% confidence intervals (CIs) were calculated from all logistic regression models.

RESULTS
Characteristics of the sample in 6th grade are provided in Table 1. Fifty percent of the children were male, 81% were white, and 18% (n = 139 of 785) were overweight. Mothers reported that children slept on average 9.0 hours per night. Overweight children in 6th grade were more likely to be male; to be of minority race/ethnicity; and to have lower quality home environments, more internalizing behaviors, and mothers with lower education. Overweight children were reported to have shorter sleep duration in 6th grade than nonoverweight children.

Relationships between each of the covariates and mother-reported sleep duration and problems at 6th grade are presented in Table 2. Boys were reported to sleep fewer hours, whereas girls were reported to have more sleep problems. Children of minority race/ethnicity were reported to have both shorter sleep duration and more sleep problems. There was a marginal but significant correlation between a higher CHAOS score and both shorter sleep duration and more sleep problems. Higher HOME scores were marginally but significantly associated with fewer sleep problems. Higher lax-parenting subscores were marginally but significantly associated with shorter sleep duration, and higher CBCL internalizing and externalizing subscale scores both were...
marginally but significantly associated with shorter sleep duration and more sleep problems.

**Adjusted Concurrent Relationship: Sleep Duration and Overweight at 6th Grade**

Greater sleep duration was associated with a reduced concurrent likelihood of overweight in 6th grade adjusted for gender, race, and maternal education (Table 3). The inclusion of bedtime significantly reduced the association between sleep duration and overweight risk (OR: 0.86; 95% CI: 0.69–1.07; *P* = .18). The inclusion of neither wake time nor general sleep problems score reduced the association between sleep duration and overweight risk. The inclusion of the CHAOS score, the HOME total score, the lax-parenting subscale score, or the CBCL internalizing or externalizing subscale scores did not alter the relationship between sleep duration and overweight risk. The CHAOS score, HOME total score, lax-parenting subscale score, and CBCL externalizing subscale score were also not independently associated with overweight in 6th grade in these models. The CBCL internalizing subscale score was, however, indepen-
shorter sleep duration in 6th grade was associated with increased concurrent risk for overweight. In addition, shorter sleep duration in 3rd grade was associated with increased likelihood of future overweight in 6th grade, independent of the child’s weight status in 3rd grade. For every additional 1 hour of sleep in 6th grade, the child was ~20% less likely (95% CI: 2%–35%) to be overweight in 6th grade. For every additional 1 hour of sleep in 3rd grade, the child was ~40% less likely (95% CI: 1%–64%) to be overweight in 6th grade. The protective effect of longer sleep duration was mediated by sleep onset rather than wake times.

To our knowledge, this study is the first to examine the association between short sleep duration and overweight risk in a relatively large US cohort. Previous studies that investigated the association between short sleep duration and overweight risk were limited by either a relatively small sample size or ethnic and socioeconomic homogeneity, making it difficult to ascertain whether the findings were easily broadly extrapolated. Our findings indicate that shorter sleep duration is associated with overweight risk in US children regardless of gender, race, or maternal education.

Children with sleeping disorders exhibit increased hyperactivity, inattention, conduct disorders, and aggression. Although in this study, externalizing behaviors alone did not seem to mediate the relationship between short sleep duration and increased overweight risk, it is possible that other factors that were not measured and that may mediate such a pathway are involved. Parents may use food to pacify sleep-deprived, irritable, and behaviorally dysregulated children. These children may also request food more often and eat beyond satiety; therefore, the effect of the sleep-deprived child’s behavior on weight may vary on the basis of parenting style. The same parents who report less control over their child’s intake may also be less strict about (or unable to control) their child’s bedtime. We had hypothesized that the relationship between short sleep duration and overweight risk may therefore have been confounded by parenting behaviors, but our data did not support this hypothesis. It is possible that family climate, including the degree of interparental or parent–child conflict, acts as a confounder, and our measure of parenting did not

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Concurrent Sleep Duration and Overweight in 6th Grade (N = 785)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Sleep duration, h</td>
<td>0.80 (0.65–0.98) a</td>
</tr>
<tr>
<td>Maternal education, y</td>
<td>0.87 (0.80–0.94) b</td>
</tr>
<tr>
<td>Gender (female vs male)</td>
<td>0.60 (0.41–0.87) a</td>
</tr>
<tr>
<td>Race (other vs white)</td>
<td>1.58 (1.02–2.46) a</td>
</tr>
</tbody>
</table>

P < .05.

<table>
<thead>
<tr>
<th>TABLE 4</th>
<th>Sleep Duration Between 3rd and 6th Grades and Overweight in 6th Grade (N = 706)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Sleep duration in 3rd grade, h</td>
<td>0.60 (0.36–0.99) a</td>
</tr>
<tr>
<td>Change in sleep duration between 3rd and 6th grades, h</td>
<td>0.68 (0.44–1.06)</td>
</tr>
<tr>
<td>Gender (female vs male)</td>
<td>0.82 (0.40–1.71)</td>
</tr>
<tr>
<td>Race (other vs white)</td>
<td>1.42 (0.54–3.73)</td>
</tr>
<tr>
<td>Maternal education, y</td>
<td>0.84 (0.72–0.99) a</td>
</tr>
<tr>
<td>BMI z score at 3rd grade</td>
<td>127.4 (48.0–337.8) b</td>
</tr>
</tbody>
</table>

P < .05.

P < .01.
assess these factors. We had also hypothesized, using the same line of reasoning, that children with more internalizing behavior problems may both have disrupted sleep and be more likely to be overweight, but the data also did not support a confounding role of internalizing behavior problems in the relationship between sleep duration and overweight risk. In short, if sleep problems and behavior problems devolve from the same underlying genetically or environmentally determined physiologic base, then our findings suggest that this common underlying mechanism is not the same mechanism that links sleep duration to overweight.

In short, the temporal relationship indicated in this study, as well as the absence of confounding by a variety of covariates, supports hypotheses set forth by others previously regarding a biological link between sleep duration and obesity risk. The proposed physiologic mechanisms have involved changes in levels of circulating leptin and ghrelin, both of which have been implicated in the regulation of appetite. Short sleep duration has also been shown to alter carbohydrate metabolism and lead to impaired glucose tolerance, which may also affect weight status. Apart from the direct influence of sleep on these parameters, there are significant endogenous circadian rhythms that affect circulating leptin, glucose, and insulin levels. The release of leptin, which reduces appetite, seems to be regulated by the circadian pace-maker, as well as increased by sleep. Reduced sleep duration has been linked with reduced leptin secretion during a 24-hour period. When sleep phase is shifted (with shift work being the most extreme example), perturbations in the stability of leptin levels would theoretically be the result, with potentially significant impacts on appetite regulation. The potential effect of delayed sleep onset on leptin secretion in children is an important area for future research.

Longer sleep duration has also been associated with a greater amount of exercise, although the direction of the association remains entirely unclear. It is possible that greater sleep duration leads to a higher likelihood that the child will be more physically active, which leads to a lower risk for overweight. Alternatively, more physical activity may both increase sleep and reduce weight gain. Notably, several previous studies of adults that tested physical activity as a covariate did not find it to be a significant contributor to the relationship between sleep and obesity risk. Ultimately, additional research is required to characterize better the physiologic components that contribute to the association between short sleep duration and overweight risk and to shed light on the preferential role of bedtime as opposed to wake time in driving the association.

In contrast to the findings regarding sleep duration, this study did not detect an association between the mother’s report of the child’s having sleep problems and likelihood of overweight in 6th grade. To the best of our knowledge, this is the first study to examine this relationship in a large sample of US children. Our findings are in agreement with studies of adults and adolescents that also did not detect an association.

There were several limitations to our study. The version of the CSHQ used in this study was modified from its original version for which psychometric data were available. The ability to compare results of the CSHQ in this study with results of other studies using the original version is therefore somewhat limited. Objective measures of sleep problems such as polysomnography were not performed, and no questions were asked about snoring or breathing difficulties that might suggest undiagnosed obstructive sleep apnea, which is known to be associated with overweight status in children. These questions deserve additional research. Parental weight status was also not available in this data set, and it is possible that this may have acted as a confounder.

Decreased sleep quantity and poor sleep quality have been associated with increased aggression, conduct disorders, impaired working memory, and poorer academic performance in children and young adolescents. It is important to note that no study to date, including this one, has established a causal relationship between sleep deprivation and increased obesity risk. Observational studies such as these can establish only that an association exists; therefore, interpreting causality should be done with caution. Nonetheless, this study suggests that an increased risk for overweight is yet another potential consequence of short sleep duration, providing an additional reason to ensure that children are receiving adequate sleep, primarily through enforcing an age-appropriate bedtime.

Despite these associations and the increasing prevalence of long-term sleep deprivation in children, the 2004 Sleep in American Poll found that only 38% of
parents with school-aged children reported that their child’s doctor asked about sleep habits. The many apparent associations between adequate sleep and optimal functioning of the child in multiple domains, including overweight status, support the importance of including a discussion of an appropriate bedtime during visits with pediatric providers. From a policy perspective, our findings also provide additional support for policies that propose later school start times. The very early school start times for US adolescents have raised concerns in the pediatric community because of their apparent adverse impact on sleep duration and, consequently, children’s general academic and behavioral functioning.63–65 The results of this study suggest that a reduced prevalence of overweight may also be a positive outcome of such policies.

APPENDIX: SLEEP PROBLEMS QUESTIONS ON CSHQ

1. My child goes to bed at the same time each night.
2. My child falls asleep within 20 minutes after going to bed.
3. My child falls asleep in own bed.
4. My child needs me or another parent in the room to fall asleep.
5. My child is afraid of sleeping alone.
6. Getting my child to bed at night is a problem.
7. My child sleeps too little.
8. My child sleeps too much.
9. My child sleeps in someone else’s bed during the night (parent, brother, sister, etc).
10. My child complains about problems sleeping.
11. My child awakens during the night screaming, sweating, and inconsolable.
12. My child awakens alarmed by a frightening dream.
13. My child awakes once during the night.
14. My child awakes more than once during the night.
15. How much of a problem are sleep wakings for you?
16. If your child wakes during the night, how much time does the night waking usually last?
17. My child wakes up in a negative mood.
18. My child has difficulty getting out of bed in the morning.
19. My child takes a long time to become alert in the morning.
20. My child wakes up very early in the morning.
21. Getting my child up in the morning is a problem.
22. My child naps during the day.
23. My child suddenly falls asleep in the middle of watching television, readings in a car, or other daily activities.
24. My child seems tired during the day.
25. Daytime sleepiness is a problem for my child.

ACKNOWLEDGMENTS
This study was supported by American Heart Association Fellow-to-Faculty Transition Award 0275040N and American Heart Association Midwest Affiliate Grant-in-Aid 0455563Z (to Dr Lumeng).

We thank Kyung Rhee, MD, Emily Fredericks, PhD, and Ron Chervin, MD, for thoughtful review of earlier versions of this manuscript.

REFERENCES
adolescents: shorter sleep, poorer sleep quality, sleepiness, and sleep-disordered breathing. J Pediatr Psychol. 2006;104:1–11
tion between reduced sleep and weight gain in women. Am J Epidemiol. 2006;164:947–954
57. Chaput JP, Despres JP, Bouchard C, Tremblay A. Short sleep duration is associated with reduced leptin levels and increased adiposity: results from the Quebec family study. Obesity. 2007;15:253–261

THE VALUE OF CASE REPORTS

“Case reports are well known to be the lowest and most useless form of evidence. Various commentators on evidence-based medicine have been known to snort derisively when they are mentioned as ‘potential papers,’ and they are dying out of most major journals. However, what is ‘well known’ is not necessarily true. Consider, for example, the diagnostic value of a raised CRP. There are some, limited instances where case reports and well observed series can obviate the need for randomized trials. Pediatricians can turn to the value of empirical antibiotics in febrile neutropenia, inhaled salbutamol in acute asthma or the classical appearance of a child with Down syndrome. These instances have something in common—they are all examples of ‘all or none’ (or ‘almost all and nearly none’) effects. Before a treatment, everyone dies. After a treatment, some don’t. It can be statistically calculated when an interesting observation becomes profound enough to be truth (Glasziou et al and e-responses). This technique compares the rate of something happening before an intervention, and the rate afterwards. If the success rate is about ten times (or more) greater with than without the intervention, then it’s probably a real effect. However, there are far more times when a single case doesn’t prove anything than occasions when it does, but don’t let anyone tell you the case report is useless.”

Arch Dis Child 2007;92:817–824
Noted by JFL
Infant-Feeding Methods and Childhood Sleep-Disordered Breathing

Hawley Evelyn Montgomery-Downs, PhD, Valerie McLaughlin Crabtree, PhD, Oscar Sans Capdevila, MD, David Gozal, MD

*Department of Psychology, West Virginia University, Morgantown, West Virginia; †Division of Pediatric Sleep Medicine, Department of Pediatrics, University of Louisville, Louisville, Kentucky

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Childhood sleep-disordered breathing has an adverse impact on cognitive development, behavior, quality of life, and use of health care resources. Early viral infections and other immune-mediated responses may contribute to development of the chronic inflammation of the upper airway and hypertrophic upper airway lymphadenoid tissues underlying childhood sleep-disordered breathing. Breastfeeding provides immunologic protection against such early exposures. Therefore, we sought to explore whether sleep-disordered breathing severity would differ for children who were breastfed as infants.

METHODS. The parents or guardians of 196 habitually snoring children (mean ± SD: 6.7 ± 2.9 years old) who were undergoing overnight polysomnography at Kosair Children’s Hospital Sleep Medicine and Apnea Center completed a retrospective survey on the method(s) used to feed the child as an infant.

RESULTS. Among habitually snoring children, those who were fed breast milk for at least 2 months had significantly reduced sleep-disordered breathing severity on every measure assessed, including apnea-hypopnea index, oxyhemoglobin desaturation nadir, and respiratory arousal index. Breastfeeding for longer than 5 months did not contribute additional benefits.

CONCLUSIONS. Our findings support the notion that breastfeeding may provide long-term protection against the severity of childhood sleep-disordered breathing. Future research should explore mechanism(s) whereby infant-feeding methods may affect the pathophysiology of development of childhood sleep-disordered breathing.
SLEEP-DISORDERED BREATHING (SDB), a condition marked by intermittent hypoxia and hypercapnia and sleep fragmentation, is present in 1 to 3% of children. Children with SDB are more frequent users of health care services, experience more frequent comorbid chronic illnesses, and may develop significant clinical cardiovascular morbidity. They also display significant behavioral and cognitive comorbidities, as well as reduced quality of life and mood alterations. Animal models of SDB display long-term, partially irreversible neurocognitive consequences after intermittent hypoxia during sleep, particularly when such exposures occur early during development.

The relatively high rate at which SDB is diagnosed among otherwise healthy children and the implications for untreated SDB are a major health concern, to the extent that the American Academy of Pediatrics recommended that primary caregivers routinely screen their pediatric patients for the presence of snoring, the cardinal symptom of SDB. The preferential treatment of obstructive sleep apnea in otherwise healthy children is surgical removal of adenoids and tonsils.

Although the pathophysiologic mechanisms underlying hypertrophy of these upper airway lymphadenoid tissues has remained unclear, direction can be taken from current studies of the lower airway suggesting that early exposure to respiratory syncytial virus infection and other frequent respiratory viral pathogens may lead to potential neurogenic inflammation and airway remodeling, both of which may contribute to development of childhood asthma. In the United States, nearly all children are infected with respiratory syncytial virus by their third year. Preliminary evidence for the presence of similar pathways in development of upper airway lymphadenoid tissue hypertrophy has recently emerged. Furthermore, inflammatory markers have been found in the adenotonsillar tissue of children who undergo surgery for SDB, and treatment with anti-inflammatory medications significantly reduces the respiratory parameters on polysomnography in these children.

Breast milk and colostrum provide the infant with immunoglobulin A and are considered a primary method for prevention of respiratory viral infections. However, the impact of breastfeeding practices on the severity of SDB in children has not been heretofore examined. We therefore sought to test the hypothesis that children who had been fed breast milk as infants would be less likely to develop significant childhood SDB compared with children who had not been fed breast milk as infants and that the duration of breastfeeding may affect such association.

METHODS

Parents of habitually snoring children who were undergoing overnight polysomnography as either a clinic patient or a participant in a larger research study at the Kosair Children’s Hospital Sleep Medicine and Apnea Center were invited to fill out a brief survey about the feeding method(s) that they used when their child was an infant. A cover letter that described implied consent and Health Insurance Portability and Accountability Act authorization was administered as approved by the institutional review boards at the University of Louisville and Norton Healthcare.

Feeding-Methods Survey

The paper-and-pencil survey included child demographics (age, gender, ethnicity, and birth order), whether anyone in the household smoked, and the relationship of the survey respondent to the child. The respondent was asked to identify the method(s) used to feed the child during the first year of life: formula only (no breast milk), breast milk only (no formula), or both formula and breast milk. For those who were fed any breast milk, the age in months at which the child was completely weaned was also identified. Finally, the respondent was asked, “How well do you remember this information?” on a 5-point scale from “not at all confident” to “completely confident.”

Polysomnography and Scoring

With the parent or guardian present, overnight polysomnography was performed by using commercially available multichannel data-acquisition equipment and scoring software (MedCare Diagnostics, Amsterdam, Netherlands) to record 8 channels of electroencephalography (O1/O2, P3/P4, C3/C4, F3/F4), submental chin and anterior tibialis electromyography, bilateral electrooculogram, snore sensor, electrocardiogram, chest and abdominal inductance plethysmography, pulse oxygen saturation (SpO2) and wave form, and oronasal airflow (assessed through an oronasal thermistor, end-tidal CO2, and nasal pressure transducer). Simultaneous video and audio monitoring was digitally recorded. No study was performed on a night when a child had an acute illness, fever, or nasal discharge. Records were scored by an analyst who was blinded to the infant-feeding–method status. Stage scoring was performed using standard criteria. Because criteria for arousals have not yet been established for children, arousals were defined as recommended by the American Sleep Disorders Association Task Force report and manually scored as spontaneous or respiratory related (occurring immediately subsequent to an apnea, hypopnea, oxyhemoglobin desaturation, or snore) and reported as indices on the basis of occurrence per hour of total sleep time (TST).

Central apneas were scored on the basis of cessation of oronasal flow and chest wall and abdominal movement; obstructive apneas were scored in the absence of oronasal airflow with continued chest wall and abdom-
inal movement, whereas decreases in oronasal flow \( \geq 50\% \) with continued effort were scored as hypopneas, provided that these events had a minimum duration of 2 breath-lengths and an associated \( \geq 4\% \) \( \text{SpO}_2 \) desaturation and/or arousal.\(^3\)\(^,\)\(^4\) Snoring was scored in the presence of a change in basal snore sensor levels that was additionally verified and annotated by the technologist via both in-room checks and microphone transmission. \( \text{SpO}_2 \) nadir and oxyhemoglobin desaturations were calculated from valid \( \text{SpO}_2 \) tracings during sleep with values during movement artifact excluded. Scoring of all respiratory events and oxyhemoglobin desaturations were initially automated and then user verified. Respiratory events were scored as indices on the basis of the number of events per hour of TST. An apnea-hypopnea index (AHI) was calculated on the basis of the number of combined apneas and hypopneas per hour of TST.

**Statistical Analyses**

Descriptive statistics were calculated. Group comparisons with dichotomous variables were made by using the \( \chi^2 \) test; those with continuous variables were performed by using 1-way analysis of variance. Data were analyzed by using SPSS 14.0 (SPSS, Chicago, IL), and \( P < .05 \) was considered statistically significant. Cohen’s \( d \) with pooled SDs was calculated to determine effect sizes.

**RESULTS**

Of the 212 surveys completed, polysomnography data were available for 196 children. There were no differences on demographic measures between habitually snoring children who were recruited from the larger research study (\( n = 89 \)) and clinical patients (\( n = 107 \)); therefore, their databases were combined. Child demographics are shown on Table 1. Surveys were completed by the child’s mother (95.9%), father (2.6%), or grandparent (1.0%). The confidence with which the respondent remembered the information was reported at 3.20 (±0.93 [SD]) on a 5-point scale, where 0 = not at all confident and 4 = completely confident. Five respondents (2.6% of sample) reported a 0 or 1 on the recall question; removal of data from these low-recall respondents did not alter any analysis, and they were retained.

When they were infants, 52% of the children were fed formula only, 10% were fed breast milk only, and 38% were fed a combination of formula and breast milk. For those who were fed breast milk in any amount, the average age of weaning was 7.3 \( \pm \) 7.0 months. No age, gender, ethnicity, or household environmental tobacco smoke exposure differences were found on feeding method, duration of breastfeeding, or polysomnography measures. Group polysomnography measures are shown in Table 2.

Two group comparisons were made. First, comparisons were made between infants who were never breastfed and those who were breastfed for increasing durations up to \( \geq 12 \) months (eg, 0 vs \( \geq 1 \) month of breastfeeding, 0 vs \( \geq 2 \) months of breastfeeding). Compared with children who were never fed breast milk, those who were fed breast milk had significantly lower AHI at every duration of breastfeeding. The breastfed children also had higher \( \text{SpO}_2 \) nadir through 11 months of breastfeeding and lower respiratory arousal indices than those who were never breastfed.

Second, comparisons were made between infants who were breastfed for increasing durations compared with those who were breastfed for less than that duration (eg, \( < 2 \) vs \( \geq 2 \) months of breastfeeding, \( < 3 \) vs \( \geq 3 \) months of breastfeeding) There was an insufficient number of children who breastfed \(< 2 \) months for comparisons. (Note that only breastfed children were included in these comparisons, and annotation of breastfeeding duration \( < 2 \) months does not include 0 months.) Longer duration of breastfeeding was associated with lower AHI through 4 months of breastfeeding, higher \( \text{SpO}_2 \) nadir through 5 months of breastfeeding, and lower respiratory arousal index through 3 months of breastfeeding. Breastfeeding beyond 5 months did not seem to provide added benefits on these measures of SDB compared with infants who had been breastfed for shorter durations.

Group means, variance, and statistical differences for comparisons between no breastfeeding and increasing durations of breastfeeding as well as between breastfeeding of different durations are shown in Fig 1 for AHI, Fig 2 for \( \text{SpO}_2 \) nadir, and Fig 3 for respiratory arousal index. There were no spontaneous arousal index differences between groups on the basis of any duration of breastfeeding. The number of children who were breastfed for

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Child Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td>Value</td>
</tr>
<tr>
<td>Age at polysomnography, mean (SD), y</td>
<td>6.7 (2.9)</td>
</tr>
<tr>
<td>Female, %</td>
<td>39.8</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White, %</td>
<td>64.3</td>
</tr>
<tr>
<td>Black, %</td>
<td>25.5</td>
</tr>
<tr>
<td>Other, %</td>
<td>10.1</td>
</tr>
<tr>
<td>Birth order, %</td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>46.9</td>
</tr>
<tr>
<td>Second or later</td>
<td>51.0</td>
</tr>
<tr>
<td>Other</td>
<td>2.0</td>
</tr>
<tr>
<td>Smoking household, %</td>
<td>30.0</td>
</tr>
</tbody>
</table>

\(^a\) Hispanic (1.5%), Asian (0.5%), bracial (5.6%), other (1.0%), and unreported (1.5%).
\(^b\) Adopted (1.0%) and unreported (1.0%).
\(^c\) Unreported (9.7%).

| TABLE 2 | Overnight Polysomnography Values |
| --- | --- | --- |
| Parameter | Group Mean (SD) | Range |
| AHI | 4.4 (8.3) | 0.0–79.1 |
| \( \text{SpO}_2 \) nadir | 87.7 (9.5) | 30.0–97.0 |
| Spontaneous arousal index | 6.5 (5.2) | 0.6–64.0 |
| Respiratory arousal index | 2.8 (5.4) | 0.0–55.7 |
successive durations and the effect sizes for statistically significant differences on each polysomnography measure comparison are shown in Table 3.

DISCUSSION

This study shows that among children who had symptoms of SDB, those who were fed breast milk when they were infants had significantly reduced disease severity. Breastfeeding for >5 months did not seem to contribute additional benefits. The differences among polysomnography measures of SDB between children who had and had not been breastfed were clinically significant: children who were breastfed had polysomnography-related severity measures below the standard threshold for surgical treatment.

We see 2 possible explanations for these findings. First, as described in the introduction, the reduced severity of SDB in breastfed children may be explained by the immunologic protection that is provided by breast milk. This notion is supported by the growing evidence suggesting that early respiratory viral exposures may contribute to the development of childhood SDB by decreasing durations; —, not applicable.

<p>| Table 3 Number of Children Breastfed for Successive Durations and Cohen’s $d$ for Significant AHI, $Sp_O_2$ Nadir, and Respiratory Arousal Index Differences |
|-------------------------------------|----------------|----------------|----------------|</p>
<table>
<thead>
<tr>
<th>Duration, Mo</th>
<th>$n$</th>
<th>$d_1$</th>
<th>$d_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>105</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>≥1</td>
<td>93</td>
<td>—</td>
<td>0.41</td>
</tr>
<tr>
<td>≥2</td>
<td>89</td>
<td>0.37</td>
<td>0.23</td>
</tr>
<tr>
<td>≥3</td>
<td>69</td>
<td>0.46</td>
<td>0.33</td>
</tr>
<tr>
<td>≥4</td>
<td>54</td>
<td>0.57</td>
<td>0.33</td>
</tr>
<tr>
<td>≥5</td>
<td>46</td>
<td>0.49</td>
<td>0.39</td>
</tr>
<tr>
<td>≥6</td>
<td>40</td>
<td>0.47</td>
<td>0.37</td>
</tr>
<tr>
<td>≥7</td>
<td>34</td>
<td>0.50</td>
<td>0.39</td>
</tr>
<tr>
<td>≥8</td>
<td>33</td>
<td>0.59</td>
<td>0.40</td>
</tr>
<tr>
<td>≥9</td>
<td>33</td>
<td>0.57</td>
<td>0.39</td>
</tr>
<tr>
<td>≥10</td>
<td>29</td>
<td>0.57</td>
<td>0.39</td>
</tr>
<tr>
<td>≥11</td>
<td>25</td>
<td>0.57</td>
<td>0.41</td>
</tr>
<tr>
<td>≥12</td>
<td>23</td>
<td>0.59</td>
<td>0.41</td>
</tr>
</tbody>
</table>

$d_1$ indicates effect size for comparison between no breastfeeding and breastfeeding for increasing durations; $d_2$, effect size for comparison between breastfeeding for increasing and decreasing durations; —, not applicable.
recurrent infections in regard to the expression and distribution of cysteinyl leukotriene receptors, suggesting that the mechanism for inflammation may be different in children with SDB.\textsuperscript{37}

Another possible explanation for our findings is that oral cavity features such as high palates, narrow dental arches, and retruded chin all are additional risk factors for SDB in children.\textsuperscript{38} Breastfeeding promotes healthy jaw formation, thereby preventing the occurrence of many of these anatomic issues\textsuperscript{39,40}; therefore, the mechanical aspects of breastfeeding may provide additional protection against development of SDB in that this feeding method promotes an upper airway that is less vulnerable to narrowing and collapse during sleep.

It is important to note that although our findings indicate that disease severity is reduced in association with breastfeeding, our work and that of others have consistently shown that even clinically mild levels of childhood SDB are associated with cognitive and behavioral compromise (see review by Beebe\textsuperscript{41}). Our work should not be interpreted to suggest that breastfeeding entirely prevents the development of SDB. Although severity was shown to be reduced in this sample, there is no consensus on what constitutes a safe level of SDB, if any, so treatment should not be delayed on the basis of a child’s feeding method.

There are several limitations to this preliminary investigation of the relationship between infant feeding and development of SDB. First, the retrospective feeding-methods survey has not been validated, although the high rate of confidence with which the respondent recalled the information was reassuring. Second, the proportion of breast milk feedings that were given by bottle is unknown and, as described, this may have some bearing on the mechanism by which breastfeeding has its beneficial effects on severity of SDB. Future investigations of this effect may use a dosage-response approach to take into account the proportion of breast milk feedings given via breast versus bottle and the type of nipple used for bottle feedings. Furthermore, although we did not find ethnicity-related differences on the incidence or duration of breastfeeding in this sample, future work should acquire more precise data to control for socioeconomic status as a possible confounder. Data on familial history of snoring and presence of SDB should also be collected in future investigations. Finally, because all of the children in this sample were habitual snorers, we did not measure differences in snoring prevalence between former breastfed and formula-fed infants. Future investigations into differences in rates and severity of SDB on the basis of infant-feeding methods is necessary.

CONCLUSIONS

Our findings support the notion that breastfeeding may provide long-term protection against the incidence and/or severity of childhood SDB. Future research aiming to explore the mechanism(s) by which infant-feeding methods may affect the pathophysiology of childhood SDB seems warranted.

ACKNOWLEDGMENTS

This study was funded by National Institutes of Health grants F32 HL-074591 (to Dr Montgomery-Downs) and HL65270 (to Dr Gozal).

We are grateful to the study participants. Jennifer Bruner, RPSGT, Nina Burkhead, Karen Conrad, Janie Cook, David Davis, Renee Ferguson, Darlene Herps, RPSGT, Contessa Howlett-Tyler, Joseph Lipovich, Tamela Nichols, Dianna O’Neal, Allison Parker, Julie Render, Andrea Shewmaker, Nigel Smith, RPSGT, Tonya Thornton-Gallahar, Nicky Wilkerson, Lisa Witcher, and Sarah Wheeler-Bayens assisted with sleep-data collection.

REFERENCES

13. Row BW, Kheirandish L, Neville JJ, Gozal D. Impaired spatial


40. Palmer B. Snoring and sleep apnoea: how it can be prevented in childhood. *Breastfeed Rev.* 2006;14:11–14
Pediatricians’ Perspectives Regarding Community Child Health: Training, Involvement, and Expectations According to Age

Cynthia S. Minkovitz, MD, MPP, Karen G. O’Connor, BS, Holly Grason, MA, Judith S. Palfrey, MD, Anita Chandra, DrPH, Thomas F. Tonniges, MD

aDepartment of Population, Family and Reproductive Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; bDivision of Health Services Research, American Academy of Pediatrics, Elk Grove Village, Illinois; cDepartment of General Pediatrics, Children’s Hospital Boston, Boston, Massachusetts; dRand Corporation, Arlington, Virginia; eBoys Town National Research Hospital, Omaha, Nebraska

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. There are increasing opportunities for pediatricians to promote children’s health through community involvement during and after residency training. Little is known about whether younger relative to established pediatricians have different experiences regarding community activities. In this study we examined whether pediatricians’ training, perspectives, and involvement in community activities vary by age.

METHODS. Eight hundred seventy-six pediatricians participated in a national, random-sample, mailed periodic survey of US members of the American Academy of Pediatrics in 2004 (response rate: 58%). χ² statistics and median tests were used to measure associations of age (≤34, 35–39, 40–50, and ≥51 years) with training, perspectives, and involvement.

RESULTS. Younger pediatricians reported more training in community child health during and before residency but were less likely to be involved currently (37.9% for ≤34 years, 44.4% for 35–39 years, 46.2% for 40–50 years, 48.3% for ≥51 years). They were more likely to report that their current involvement was too little versus just right or too much (81.3%, 73.5%, 60.7%, and 47.1%, respectively). Younger pediatricians were more willing to spend ≥1 hour/month on community child health activities (95.0%, 91.2%, 89.7%, and 85.4%, respectively). Younger versus older pediatricians were more likely to sense moderate or greater responsibility for improving children’s health in their community (83.6%, 77.2%, 76.7%, and 70.2%, respectively) and expected their community work to increase during the next 5 years (80.0%, 67.5%, 59.7%, and 40.1%, respectively). Age findings persisted when adjusted for gender.

CONCLUSIONS. Although practice constraints may limit community involvement, younger pediatricians anticipated growing participation in community activities. Longitudinal studies are needed to determine whether such expectations are realized.
Many pediatric residency programs now provide didactic and experiential training in community pediatrics. Solomon et al. reported that in 2002, 28% of programs provided training in 0 to 3 community settings, 41% in 4 to 6 settings, and 31% in ≥7 activities. In addition, many provided didactic education on various community health topics and exposed residents to communicating with elected officials, participating on longitudinal projects, conducting research in the community, and, to a lesser degree, providing legislative testimony.

Although the pediatric Residency Review Committee mandated training in child advocacy in 1991, core competencies in community pediatrics have not been adopted universally. Wright et al. used a modified Delphi process to develop a set of objectives for a resident advocacy curriculum. Rezet et al. identified community pediatrics competencies and learning objectives, cross-referenced to the Accreditation Council for Graduate Medical Education competencies, that residents are expected to achieve. Despite lack of consensus on curricular content, most residency programs offer related community pediatrics training; however, it is unclear whether pediatricians who were trained more recently have more skills or more favorable perspectives than more established pediatricians.

The objective of this study was to examine whether pediatricians’ training, perspectives, and involvement in community activities vary by age. We hypothesized that younger physicians relative to older physicians would report more training during residency, less current involvement, and comparable senses of responsibility for improving child health in their communities. These assumptions are based on the growing emphasis of community pediatrics in residency training, conflicting demands of family responsibilities and of establishing practices and/or completing training in the early years after completing residency, and a shared sense of responsibility for children’s health, regardless of age.

METHODS

Periodic Survey

The American Academy of Pediatrics (AAP) conducts periodic surveys (PSs) 3 to 4 times per year on topics of importance to pediatricians. Each survey uses a unique random sample of members of the AAP. The 2004 PS (PS 60) included questions on involvement in community child health, including a global question asking participants to indicate whether in the past 12 months they “participated in a professional capacity in any activities that promote child health” in their community. The surveys also included a separate question inquiring about involvement in 19 individual activities (and other category) and whether participation was volunteer or paid. All participants further identified their use of 8 strategies to influence children’s health in their home and practice communities, including serving on community boards; participating on child health committees; working with a coalition; educating legislators; communicating with media; addressing parent, teacher, or other community groups; working with the local AAP chapter; and voting in state/local elections. All respondents additionally reported on their perspectives regarding community involvement and their use in the past 12 months of 6 skills related to community health activities: locating community resources for individual children, identifying community needs, using population-level data to understand the determinants of children’s health, working as a member of an interdisciplinary team to promote children’s health in the community, speaking publicly on behalf of children’s health, and using computers/Internet to find information about child health policy and related activities. Respondents also described the timing of their participation and training in community activities (eg, before medical school, during medical school, during residency, during fellowship training, since completing training).

Survey content was informed by a national advisory group with expertise in community pediatrics, by review by members of the AAP Community Pediatrics Action Group and the Council on Community Pediatrics, and review of similar PS questions asked in 1989 and 1993. This survey was an 8-page self-administered questionnaire sent to 1829 active members. The original mailing and 5 follow-up mailings to nonrespondents were conducted from April through October 2004. After the first and fifth mailings, an e-mail reminder was sent to nonrespondents with e-mail addresses, and a postcard reminder was sent to those without (68% and 32% of nonrespondents, respectively). We received a total of 1053 completed questionnaires for a response rate of 57.6%.

Data Analysis

Data analysis was conducted to examine the relationship between age and involvement in community child health activities. Analysis included postresidency pediatricians and excluded residents (n = 160) and pediatricians with a Specialty Fellow designation (n = 12) in the AAP membership database as well as 5 respondents who did not provide age information. The final sample included 876 pediatricians (83.2% of respondents).

χ² statistics were used to assess differences according to age in demographic and practice characteristics, community child health perspectives, and participation in community child health activities. Median test analyses (Kruskal-Wallis) also were used to compare skill level in community child health activities according to age category. Age analyses were adjusted for the effect of gender. For the purposes of analysis, 4 age categories were constructed: ≤34, 35 to 39, 40 to 50, and ≥51 years). The age range for the youngest category (≤34 years) was identified, as a proxy, to capture the experiences of pediatricians within their first 5 years out of residency.
This is consistent with graduating residents’ mean age being 32 years and 88% of residents in the United States completing residency training at 34 years (K.G.O., unpublished data). In general, 5 to 10 years are required after residency to become established in pediatric practice and to develop relationships in the community.7

For assessing potential response bias in 2004, comparisons between respondents and nonrespondents were conducted for several demographic variables. No significant differences were found between respondents and nonrespondents for mean age (43.7 years) and region of country (24.5% Northeast, 21.5% Midwest, 33.4% South, and 20.7% West). More respondents were female (53.9% vs 46.6%; P < .05).

Analyses were conducted by using SPSS 11.5 (SPSS Inc, Chicago, IL). Human-subjects approval was obtained from the AAP Institutional Review Board and the Committee on Human Research at Johns Hopkins Bloomberg School of Public Health.

RESULTS

Respondent Characteristics

Demographic and practice characteristics of respondents were compared according to age (Table 1). A greater
TABLE 2  Involvement in Community Child Health Activities in Past 12 Months According to Age

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Total (N = 876), n (%)</th>
<th>Age, n (%), y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;=34 (n = 141)</td>
<td>35–39 (n = 164)</td>
</tr>
<tr>
<td>Participate in community activities to promote child health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>385 (45.1)</td>
<td>53 (37.9)</td>
</tr>
<tr>
<td>No</td>
<td>468 (54.9)</td>
<td>87 (62.1)</td>
</tr>
<tr>
<td>Payment in community activitiesa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not paid (volunteer)</td>
<td>306 (79.7)</td>
<td>46 (86.8)</td>
</tr>
</tbody>
</table>

a Among subset of respondents who reported community child health activity participation (valid percentages reported); P < .05.

percentage of respondents who were &le;50 years were female, whereas more male respondents were represented among those who were &ge;51 years. There were no age differences in racial/ethnic distribution. Approximately half of respondents were working in pediatric group or health maintenance organization practices, and more than half were located in urban settings. Older pediatricians were more commonly in solo or 2-physician practices compared with young pediatricians. More than one quarter of all respondents spent less than half of their time in general pediatrics, and pediatricians in the youngest (through 34 years) and oldest (&ge;51 years) age categories more frequently indicated this time distribution.

### Participation in Community Child Health

Table 2 highlights involvement in community child health activities by the 4 age categories. Participation in community child health tended to be least frequent in the youngest age group (37.9% for &le;34 years, 44.4% for 35–39 years, 46.2% for 40–50 years, and 48.3% for &ge;51 years; P = .51). The majority of respondents indicated that their community involvement is only on a volunteer basis, with the youngest begin most likely to volunteer (86.8% for &le;34 years, 79.7% for 35–39 years, 85.1% for 40–50 years, and 71.1% for &ge;51 years; P < .05).

There were age differences in the timing of participation in community child health activities during the course of medical training. More young pediatricians compared with older pediatricians indicated involvement in the early part of their career, including before (68.2%, 59.3%, 43.0%, and 26.9%, respectively; P < .01) and during medical school (80.3%, 71.1%, 46.5%, and 36.8%, respectively; P < .01). In addition, a greater percentage of young pediatricians participated in community child health training in residency programs (74.2% for &le;34 years, 58.5% for 35–39 years, 51.3% for 40–50 years, and 38.3% for &ge;51 years; P < .001); however, pediatricians in the older age groups reported more involvement since completing residency than young pediatricians (36.4% for &le;34 years, 56.3% for 35–39 years).

### TABLE 3  Perspectives on Involvement and Responsibility for Child Health According to Age

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Total (N = 876), n (%)</th>
<th>Age, n (%), y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;=34 (n = 141)</td>
<td>35–39 (n = 164)</td>
</tr>
<tr>
<td>Current level of involvement in community child health activitiesa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Too little</td>
<td>536 (62.1)</td>
<td>113 (81.3)</td>
</tr>
<tr>
<td>Just right</td>
<td>324 (37.5)</td>
<td>26 (18.7)</td>
</tr>
<tr>
<td>Too much</td>
<td>3 (0.3)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>View of child health responsibilityb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very responsible</td>
<td>283 (32.8)</td>
<td>40 (35.0)</td>
</tr>
<tr>
<td>Moderately responsible</td>
<td>371 (43.0)</td>
<td>68 (48.6)</td>
</tr>
<tr>
<td>A little responsible</td>
<td>195 (22.6)</td>
<td>23 (16.4)</td>
</tr>
<tr>
<td>Not at all responsible</td>
<td>13 (1.5)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Time willing to spend in child health activities, h/moa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;5</td>
<td>93 (10.8)</td>
<td>10 (7.1)</td>
</tr>
<tr>
<td>4–5</td>
<td>84 (9.8)</td>
<td>20 (14.2)</td>
</tr>
<tr>
<td>1–3</td>
<td>435 (50.6)</td>
<td>89 (63.1)</td>
</tr>
<tr>
<td>&lt;1</td>
<td>158 (18.4)</td>
<td>15 (10.6)</td>
</tr>
<tr>
<td>0</td>
<td>90 (10.5)</td>
<td>7 (5.0)</td>
</tr>
<tr>
<td>Expectation of community work in next 5 ya a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase</td>
<td>501 (58.3)</td>
<td>112 (80.0)</td>
</tr>
<tr>
<td>Stay the same</td>
<td>336 (39.1)</td>
<td>28 (20.0)</td>
</tr>
<tr>
<td>Decrease</td>
<td>22 (2.6)</td>
<td>0 (0.0)</td>
</tr>
</tbody>
</table>

a P < .01.
b P < .05.
years, 78.1% for 40–50 years, and 87.6% for ≥51 years). Age differences in the timing of training persisted when analyses were adjusted for gender.

Pediatricians chose varying strategies in their community child health promotion activities. Overall, most respondents used voting in state and local elections (69.7%) and addressing community groups (48.7%) as their more common strategies rather than working as a coalition member (8.4%) or serving on a community organization board (17.0%). Pediatricians in the oldest and youngest age groups educated legislators (26.9% for ≤34 years, 11.5% for 35–39 years, 19.9% for 40–50 years, and 33.3% for ≥51 years; P < .01) and worked with their local AAP chapters (21.2% for ≤34 years, 10.7% for 35–39 years, 14.5% for 40–50 years, and 23.1% for ≥51 years; P < .01). Conversely, more pediatricians between 35 and 50 years of age focused on addressing parent, teacher, or other community groups (46.2% for ≤34 years, 59.5% for 35–39 years, 52.0% for 40–50 years, and 40.0% for ≥51 years; P < .01).

The community child health settings where pediatricians were engaged (health, school, government, or nonprofit) did not vary by age, with 1 exception. Older pediatricians were more involved with nonprofit organizations (eg, Children’s Defense Fund, March of Dimes) than younger pediatricians (11.3% for ≤34 years, 11.6% for 35–39 years, 16.3% for 40–50 years, and 20.7% for ≥51 years; P < .05).

Community Child Health Perspectives
There were key age differences in perspectives toward involvement in community child health activities (Table 3). A greater percentage of young pediatricians reported that their current level of participation is too little, whereas more than half of pediatricians in the oldest age category (≥51 years) believed that their involvement was “just right.” Young pediatricians also expected an increase in their community work in the next 5 years, whereas fewer pediatricians who were older than 40 years anticipated an increase (81.3% for ≤34 years, 73.5% for 35–39 years, 60.7% for 40–50 years, and 47.1% for ≥51 years; P < .01). Younger versus older pediatricians were more likely to sense moderate or greater responsibility for improving the health of children in their community (83.6% for ≤34 years, 77.2% for 35–39 years, 76.7% for 40–50 years, and 70.2% for ≥51 years; P < .05). Age differences persisted after adjustment for gender.

Skills and Training in Community Child Health
Tables 4 and 5 describe use of 6 skills in participating in community health activities and reported skill levels ac-

### TABLE 4
Skills Used in Participating in Community Child Health Activities in the Past Year According to Age

<table>
<thead>
<tr>
<th>Activity</th>
<th>Total (N = 876), n (%)</th>
<th>≤34 (n = 141)</th>
<th>35–39 (n = 164)</th>
<th>40–50 (n = 295)</th>
<th>≥51 (n = 276)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locate resources for individual children&lt;sup&gt;a&lt;/sup&gt;</td>
<td>564 (68.4)</td>
<td>96 (69.1)</td>
<td>114 (73.5)</td>
<td>191 (68.2)</td>
<td>163 (64.9)</td>
</tr>
<tr>
<td>Identify community needs&lt;sup&gt;b&lt;/sup&gt;</td>
<td>248 (30.8)</td>
<td>26 (19.4)</td>
<td>42 (28.2)</td>
<td>85 (31.3)</td>
<td>95 (37.8)</td>
</tr>
<tr>
<td>Use population-level data to understand the determinants of children’s health</td>
<td>201 (25.2)</td>
<td>32 (24.1)</td>
<td>27 (18.0)</td>
<td>68 (25.0)</td>
<td>74 (30.5)</td>
</tr>
<tr>
<td>Member of a team to promote child health</td>
<td>246 (30.4)</td>
<td>34 (25.2)</td>
<td>38 (25.5)</td>
<td>89 (32.2)</td>
<td>85 (34.0)</td>
</tr>
<tr>
<td>Speak publicly on behalf of children’s health&lt;sup&gt;b&lt;/sup&gt;</td>
<td>227 (27.8)</td>
<td>31 (22.8)</td>
<td>33 (21.6)</td>
<td>85 (30.9)</td>
<td>78 (31.0)</td>
</tr>
<tr>
<td>Use computers and Internet to find information about child health policy and related activities&lt;sup&gt;a&lt;/sup&gt;</td>
<td>550 (67.0)</td>
<td>109 (79.0)</td>
<td>108 (70.6)</td>
<td>180 (65.0)</td>
<td>153 (60.5)</td>
</tr>
</tbody>
</table>

<sup>a</sup> P < .01.  
<sup>b</sup> P < .05.

### TABLE 5
Moderate/High Skills Used in Participating in Community Child Health Activities According to Age

<table>
<thead>
<tr>
<th>Activity</th>
<th>Total (N = 876), n (%)</th>
<th>≤34 (n = 141)</th>
<th>35–39 (n = 164)</th>
<th>40–50 (n = 295)</th>
<th>≥51 (n = 276)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locate resources for individual children&lt;sup&gt;a&lt;/sup&gt;</td>
<td>397 (56.6)</td>
<td>51 (44.0)</td>
<td>73 (53.3)</td>
<td>138 (56.8)</td>
<td>135 (65.5)</td>
</tr>
<tr>
<td>Identify community needs&lt;sup&gt;b&lt;/sup&gt;</td>
<td>167 (23.7)</td>
<td>14 (17.5)</td>
<td>24 (23.1)</td>
<td>62 (31.5)</td>
<td>67 (37.0)</td>
</tr>
<tr>
<td>Use population-level data to understand the determinants of children’s health</td>
<td>156 (28.1)</td>
<td>20 (23.5)</td>
<td>24 (23.3)</td>
<td>52 (26.3)</td>
<td>60 (35.3)</td>
</tr>
<tr>
<td>Member of a team to promote child health</td>
<td>253 (43.9)</td>
<td>26 (30.6)</td>
<td>48 (44.9)</td>
<td>84 (41.0)</td>
<td>95 (53.1)</td>
</tr>
<tr>
<td>Speak publicly on behalf of children’s health&lt;sup&gt;b&lt;/sup&gt;</td>
<td>293 (50.7)</td>
<td>35 (41.7)</td>
<td>44 (41.5)</td>
<td>117 (56.0)</td>
<td>97 (54.2)</td>
</tr>
<tr>
<td>Use computers and Internet to find information about child health policy and related activities&lt;sup&gt;a&lt;/sup&gt;</td>
<td>421 (60.2)</td>
<td>85 (70.2)</td>
<td>89 (66.9)</td>
<td>141 (60.3)</td>
<td>106 (50.2)</td>
</tr>
</tbody>
</table>

Significance testing using Kruskal-Wallis test for medians.  
<sup>a</sup> P < .01.  
<sup>b</sup> P < .05.
ccording to age. Using the Internet to find information about child health policy was more common among younger pediatricians, whereas working to identify community health needs was cited less commonly. Median test analyses revealed that younger pediatricians reported a higher skill level in using the computers/Internet to find information about child health policy \((P < .01)\), whereas older pediatricians noted more skill in locating resources for children \((P < .01)\), speaking publicly on behalf of child health \((P < .05)\), and identifying community needs \((P < .05)\). Age differences in skills with regard to computers, local resources, and identifying community needs persisted when adjusted for gender.

Training in community child health also differed according to age. Many (44%) respondents had never received formal training at any point in their medical career, and the most common time for training was during residency (39.1%). Formal training during medical school (43.0%) for \(\leq 34\) years, 52.6% for 35–39 years, 32.0% for 40–50 years, and 22.4% for \(\geq 51\) years; \(P < .01\) and residency (63.4% for \(\leq 34\) years, 54.7% for 35–39 years, 29.9% for 40–50 years, and 20.0% for \(\geq 51\) years; \(P < .01\)) was more frequently cited among younger pediatricians, whereas older pediatricians indicated receipt of training after residency (5.4% for \(\leq 34\) years, 7.4% for 35–39 years, 25.9% for 40–50 years, and 36.0% for \(\geq 51\) years; \(P < .01\)). Again, findings of differences in training according to age were independent of gender.

**DISCUSSION**

Traditionally, work in the community and particularly work that is aimed at policy change at a systems level generally has been done by more established physicians, based on their years of experience in the community and their increasing understanding of the complex interplay of systems, community, and family factors on the lives of children. In recent years, there has been a growing interest among young physicians and among the faculty in residency training programs to introduce the knowledge, skills, and attitudes necessary to address community- and systems-level problems earlier in the careers of physicians. A driving force for the creation of the Dyson Community Pediatrics Initiative was the notion that advocacy in pediatrics at the community level would be far more effective if it were an integral part of a pediatrician’s work throughout his or her career than if it were engaged in only at the culmination of the career.

In these analyses, pediatricians reported more training during medical school and residency, greater expectations for increased future involvement, and less involvement in the past year by the youngest pediatricians. The observed differences in perspectives regarding and engagement in community activities suggest a need to develop age-specific strategies for encouraging increased participation in community child health. For example, younger pediatricians may benefit from senior colleagues’ introducing them to community leaders, highlighting local resources, identifying community needs, and providing networking opportunities. More established pediatricians may benefit from training in specific skills to which they have not been previously exposed, such as computer-related activities and the additional human resources of younger pediatricians who also are committed to improving the lives of children.

Although all pediatricians balance work and personal commitments with community involvement, opportunities for younger pediatricians may need to recognize places and organizations with which young parents come into contact to facilitate their involvement in ways that are particularly meaningful to the needs of their own families. Organizations that are interested in promoting involvement of those in the middle age group may find that a mix of strategies is needed. For all groups, it is likely that pediatricians will continue to select multiple venues in which to be engaged, based on life experiences, personal preferences, and political climate.

Differences observed according to age also may reflect generational differences with those in the youngest age group, a product of generation X, more concerned about work-life balance and perhaps, in combination with challenges of establishing careers and families, being less willing to engage at the current time in community activities. More established pediatricians, born during the baby boomer and traditional eras, may be more willing to work longer hours without enhanced remuneration and be more service oriented. To address this possibility, we assessed participation in community activities among those aged 40 to 50 years in this survey, aged 29 to 39 years in the 1993 survey, and aged 29 to 35 years in the 1989 survey and found comparable percentages of 42%, 49%, and 46%, respectively, suggesting stable levels of involvement among the cohort over time; however, within each PS, we also observed increasing involvement in community activities with increasing age, suggesting that generational differences alone do not account for the observed findings; rather, younger pediatricians likely require time to establish themselves professionally and gain credibility in their local communities before engaging in community activities. In both the 2004 and 1993 surveys, greater percentages of younger pediatricians reported that their current level of involvement was too little, recognizing a desire for increasing involvement.

Several limitations should be noted. First, self-report may lead to overestimations of community involvement, although we have no reason to suspect that this varies by age. It is possible that younger pediatricians, with more training in residency and not having entered the work...
world beyond training, inflate their anticipated future involvement. Second, enhanced exposure to community pediatrics may increase social desirability bias, with younger pediatricians reporting but not truly believing that they have greater responsibility than their older colleagues. Third, we focused on population-based rather than individual patient-level skills in this national survey of pediatricians; however, other assessment tools assess quality of care with regard to medical home attributes and primary care. Fourth, it is possible that older pediatricians do not report receipt of formal training during residency because of recall bias; however, our finding of increased training for those who more recently completed residency is consistent with increased exposures to community pediatrics reported by third-year residents surveyed from 1997 to 2002. Finally, the survey response rate was 58%, although this level of participation is consistent with other national surveys of physicians, and respondent bias has not been observed in other PSs.

Whether generational perspectives ultimately influence pediatricians’ involvement in community activities is unclear. From 1997 to 2002, third-year pediatric residents reported enhanced quality of training regarding child advocacy and assessing the needs of their community as they increasingly were exposed to community sites, including schools, child care centers, public health departments, and community health centers. It is possible that increased skills and exposures will overcome potential generational tendencies for less involvement; however, with younger physicians increasingly working in salaried positions in group practices, ability to participate in community activities may require engagement outside work responsibilities if such activities are not supported and valued by employers or paid for by other sources. Although there is decreased involvement by younger pediatricians, this largely reflects a decline in paid activities. A separate article (C.S.M., K.G.O., H.G., A.C., C. A. Aligne, MD, MPH, M. D. Kogan, PhD, and D. Tayloe, MD, unpublished data) discussed time trends with regard to community child health activities and documented that during the past 2 decades, there was a decline in pediatrician involvement in government and philanthropically sponsored paid opportunities for community activities.

CONCLUSIONS
Pediatricians’ training, perspectives, and involvement in community child health activities vary by age. Although younger pediatricians report less current involvement, they also report receiving more training during and before residency and believe that their current level of involvement is insufficient. Prospective longitudinal studies, such as that being conducted as part of the national evaluation of the Dyson Community Pediatrics Training Initiative, are needed to address whether greater exposures to community training during residency and expectations for greater involvement translate into enhanced involvement once pediatricians are established in their careers.

ACKNOWLEDGMENTS
This study was funded by the Maternal and Child Health Bureau, Health Resources and Services Administration, and supported by the American Academy of Pediatrics.

These analyses were conducted as part of the Dyson Initiative National Evaluation and informed by the AAP PS 60 Workgroup: C. Andrew Aligne, MD, MPH (University of Rochester School of Medicine and Dentistry, Rochester, NY); Holly Grason, MA (Johns Hopkins Bloomberg School of Public Health, Baltimore, MD); David E. Heppel, MD (Maternal and Child Health Bureau/Health Resources and Services Administration, Rockville, MD); Michael D. Kogan, PhD (Maternal and Child Health Bureau/Health Resources and Services Administration, Rockville, MD); Cynthia S. Minkovitz, MD, MPP (Johns Hopkins Bloomberg School of Public Health, Baltimore, MD); Karen G. O’Connor (American Academy of Pediatrics, Elk Grove Village, IL); Judith S. Palfrey, MD (Children’s Hospital, Boston, MA); David Tayloe, MD (Goldsboro Pediatrics, Goldsboro, NC), and Thomas F. Tonniges, MD (Boys Town National Research Hospital, Omaha, Nebraska).

We gratefully acknowledge thoughtful review by Lynn Olson, PhD.

REFERENCES
8. Grason H, Aliza B, Hutchins VL, Guyer B, Minkovitz C. Pedi-
atician-led community child health initiatives: case summaries from the evaluation of the community access to child health program. *Pediatrics*. 1999;103(6 pt 3):1394–1419


THE STATE CHILDREN’S HEALTH INSURANCE PLAN AND THE WAR ON (EXPENSIVE) DRUGS

“The reauthorization of the State Children’s Health Insurance Program (SCHIP), created in 1997 to cover children from lower-income families who make too much to qualify for Medicaid, is up for renewal this fall. Tucked into page 414, section 904 of the House bill is a provision to spend more than $300 million to establish a new federal ‘Center for Comparative Effectiveness’ to conduct government-run studies of the economic considerations that go into drug choices. The center will initially be funded through Medicare but will soon get its own ‘trust fund.’ The aim is to arm government actuaries with data that proponents hope will provide ‘scientific’ proof that expensive new drugs are no better than their older alternatives. The trick is to maintain just enough credibility around the conduct of these trials to justify unpopular decisions not to pay for newer medicines. While there’s nothing inherently wrong with this sort of fiscally minded clinical research, Medicare is no ordinary payer: It dictates decisions made in the private market. So as the government begins tying its own payment decisions to the results of its own studies, there’s a great temptation to selectively interpret data and arbitrarily release results. Clearly, this obvious conflict of interest demands even more outside scrutiny and transparency than has been the usual fare when it comes to government research. The inherent complexity and limitations of conducting these sorts of ‘comparative’ drug trials also need to be carefully considered before policy makers rush to tie sweeping payment choices to results of single studies. If not, there’s a real risk that faux science and limited findings will be used to set rigid payment policies that will arbitrate access to new treatments for the entire health-care market.”


Editor’s Note: At first this looks like a good idea, but later you realize it can be a very bad idea.

Noted by JFL
Health Status and Health-Related Quality of Life Preference-Based Outcomes of Children Who Are Aged 7 to 9 Years and Have Bilateral Permanent Childhood Hearing Impairment

Stavros Petrou, PhD¹, Donna McCann, PhD², Catherine M. Law, MD³, Peter M. Watkin, MB, BS, MSc⁴, Sarah Worsfold, BScSc⁵, Colin R. Kennedy, MD⁶

¹National Perinatal Epidemiology Unit and Health Economics Research Centre, University of Oxford, Oxford, United Kingdom; ²Department of Child Health, University of Southampton, Southampton, United Kingdom; ³Institute of Child Health, University College London, London, United Kingdom; ⁴Department of Audiology, Whipps Cross University Hospital, London, United Kingdom

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The objective of this study was to describe the health status and health-related quality of life preference-based outcomes of children with diagnosed bilateral permanent childhood hearing impairment and a comparison group of English-speaking children with normal hearing.

METHODS. We studied 120 children who were aged 7 to 9 years and had bilateral permanent childhood hearing impairment of moderate or greater severity, identified from a cohort of 156,733 children who were born in 8 districts of southern England, and 63 English-speaking children with normal hearing and the same place of birth and age at assessment. Principal caregivers were interviewed by using the Health Utilities Index Mark III questionnaire for proxy-assessed usual health status assessment. Levels of function within each of the 8 attributes of the Health Utilities Index Mark III (cognition, vision, hearing, speech, ambulation, dexterity, emotion, and pain) were recorded.

RESULTS. Bilateral permanent childhood hearing impairment is associated with significantly increased proportions of suboptimal levels of function and significantly lower single-attribute utility scores in 6 of the 8 attributes of the Health Utilities Index Mark III: vision, hearing, speech, ambulation, dexterity, and cognition. Compared with the children with normal hearing, the mean multiattribute utility score for the children with hearing impairment was significantly lower for both the whole group and the moderate, severe, and profound severity subgroups. The differences in the distributions of the multiattribute utility scores between the children with hearing impairment as a group and the children with normal hearing and between each of the severity subgroups and the children with normal hearing all were statistically significant.

CONCLUSIONS. This study provides rigorous evidence of an association between bilateral permanent childhood hearing impairment and diminished health status and health-related quality of life preference-based outcomes during midchildhood.
BILATERAL PERMANENT CHILDHOOD hearing impairment of moderate or greater severity in the early years affects approximately 1 in 750 children, and >80% is present at birth.1–3 The consequences of the condition may include lifelong impairment in language skills and possible delays in social development and academic achievement.2–5 The severity of the disability that is experienced by the child is influenced by the degree and duration of hearing loss, the age at which the hearing loss first appears, and the hearing frequencies affected.6 The severity of disability is also influenced by coexisting disabilities that the child has, such as visual impairment, learning difficulties, and cerebral palsy.6 Developmental delays are apparent for children with severe and profound hearing, as well as for those who have moderate and less severe impairment.7–9

Most studies that describe the long-term outcomes of children with hearing impairment focus on speech and language impairment and fail to capture all functional, neurodevelopmental, and behavioral outcomes that might be of interest.2 In recent years, a number of investigators have recognized the importance of measuring the impact of childhood conditions across multiple domains. Instruments that can be used to measure the multiple health impacts of bilateral permanent childhood hearing impairment include multidimensional health profiles, which measure aspects of physical, mental, and social well-being across several domains. Instruments that can be used to measure the multiple health impacts of bilateral permanent childhood hearing impairment include multidimensional health profiles, which measure aspects of physical, mental, and social well-being across several domains. The multiattribute utility measures that have been developed include the Quality of Well-being Scale,11 Rosser-Kind Classification of Illness States,12 Health Utilities Index,13 EQ-5D,14 16D,15 17D,16 Assessment of Quality of Life instrument,17 and SF-6D.18 The Health Utilities Index is the most widely used of the multiattribute utility measures within the childhood context. Three studies have used the Health Utilities Index as the basis for estimating preference-based health outcomes in economic evaluations of pediatric cochlear implantation.19–21 The objective of this study was to use the Health Utilities Index to describe the health status and health-related quality of life preference-based outcomes of children who were aged 7 to 9 years and had diagnosed bilateral permanent childhood hearing impairment of moderate or greater severity and a comparison group of English-speaking children with normal hearing.

METHODS

Study Background and Sample

This study was undertaken as part of a larger program of research, the Hearing Outcomes Project, which focused on the speech and oral language abilities of children with bilateral permanent childhood hearing impairment.22 The study sample included all children with bilateral permanent childhood hearing impairment of at least 40 dB hearing level (HL) identified from a cohort of 156 733 children who were born in 8 districts of southern England: Southampton, Portsmouth, Swindon, Bath, Waltham Forest, Hillingdon, Redbridge, and Brent and Harrow. The districts broadly reflected the nation’s socioeconomic characteristics with unemployment rates in 2001 in Southampton, Portsmouth, Swindon, Bath, Waltham Forest, Hillingdon, Redbridge, and Brent and Harrow of 2.88%, 3.09%, 2.46%, 1.98%, 4.92%, 2.73%, 3.64%, and 4.14%, respectively, compared with a national average of 3.35%.23 Similarly, the districts broadly reflected the nation’s underlying health characteristics with the proportion of the population with a limiting long-term illness in Southampton, Portsmouth, Swindon, Bath, Waltham Forest, Hillingdon, Redbridge, and Brent and Harrow estimated at 17.43%, 17.44%, 15.26%, 15.85%, 16.54%, 14.88%, 16.30%, and 15.28%, respectively, compared with a national average of 17.93%.23 The severity of hearing impairment, age at confirmation, and age at institution of management of cases of bilateral permanent childhood hearing impairment were obtained by review of the case records of pediatric audiologists, audiology scientists, family practitioners, speech and language therapists, and other professionals involved in each district. Children with a known postnatal cause of bilateral permanent childhood hearing impairment (eg, bacterial meningitis) were not included in the study sample. Four of the districts in which children were identified (Southampton, Portsmouth, Swindon, and Bath) are in the Wessex region and provided the sample for the years 1993 to 1996 for the Wessex controlled trial of universal newborn screening (UNS) for permanent childhood hearing impairment.24,25 The other 4 districts in which children were identified are in the Greater London region. During the period 1992 to 1997, 2 of these districts (Waltham Forest and Hillingdon) had operated a UNS program and their 2 neighboring districts (Redbridge, and Brent and Harrow) had not operated a UNS program. Both of the London districts that were operating UNS programs had reported on the impact of UNS on early diagnosis of permanent childhood hearing impairment and collected
data prospectively in this birth cohort. In the case of Waltham Forest, this had been a prospective comparison with the Redbridge district. Thus, approximately one half of the entire birth cohort of this study had been in a target population for UNS for permanent childhood hearing impairment with >90% of the sample included from birth in prospective studies of permanent childhood hearing impairment.

At the age of 7 to 9 years, the study children were traced; the principal caregiver was then contacted by mail, regardless of the current place of residence, and permission was requested to participate in this study. When consent was provided, 2 researchers who were unaware of the child’s early hearing or audiologic history arranged a time convenient for a visit to the child’s home. One researcher interviewed the principal caregiver, who was usually the mother, which included completion of a battery of research instruments. The child was simultaneously assessed by the other researcher in a separate room with the following: Test of Receptive Grammar; British Picture Vocabulary Scale (receptive language); Renfrew Bus Story Test (expressive language); and Raven’s Progressive Matrices Test (nonverbal abilities). For the purpose of comparison with a group of children without permanent bilateral permanent childhood hearing impairment, a group of English-speaking children with normal hearing were also identified and assessed using the same measures. The children with normal hearing were identified using an algorithm developed by the investigators that calculated in an incremental manner the age distribution of the children who had hearing impairment and were recruited in each district during 4-month periods. The algorithm calculated the dates of birth of the required number of children with normal hearing within each district for each 4-month period. Each district child health officer was then asked to select randomly from routine data collection systems children who were born in the target hospitals on specified dates and approach their principal caregiver for participation in the study (with a target of 4 principal caregivers to be approached for each child with hearing impairment). This process was designed to recruit a group of children who had normal hearing and similar place of birth and age at assessment to the children with hearing impairment. Full details on the tracing and recruitment procedures of the children with hearing impairment and children with normal hearing and the measures used to assess outcomes are reported elsewhere.

Measurement of Health Status and Health-Related Quality of Life Preference-Based Outcomes

As part of the home visit, the principal caregiver was interviewed about the health status and health-related quality of life of the child using the Health Utilities Index. The principal caregiver was considered the appropriate subject for the task because pilot research had indicated that the comprehension level for the Health Utilities Index is somewhat high for a pediatric sample in which a number of children may have developmental disabilities.

The Health Utilities Index is a family of preference-based, multiattribute utility measures. The principal caregiver was interviewed using the unedited 15-item questionnaire for proxy-assessed usual health status assessment, which was obtained from the Health Utilities Index developers and covers both Mark II and Mark III health status classification systems. The “usual” health focus of the questions has previously been applied in population health surveys, in which short-term illnesses such as the flu are not the major concern. The Mark III classification system is now recommended by the developers because of its broad applicability in both clinical and general population health studies, improvements in a number of definitions, and an increased orthogonality of its attributes for structural independence. The Health Utilities Index Mark III health status classification system were converted into single-attribute and multiplicative multiattribute utility scores using published utility functions. Single-attribute utility scores for the Health Utilities Index Mark III are defined on a scale from 0.00 (reflecting the preference score for the lowest level of function for the attribute) to 1.00 (reflecting the preference score for the highest level of function for the attribute). Multiattribute utility scores for the Health Utilities Index Mark III, in contrast, are based on the permutation of responses across the 8 attributes and are expressed on an interval scale ranging from −0.36 (representing the health state with the lowest level of function for all attributes) to 1.00 (representing the health state with the highest level of function for all attributes). The multiattribute utility scoring algorithm for the Health Utilities Index Mark III can be summarized as

$$u^* = 1.371(b_1 \times b_2 \times b_3 \times b_4 \times b_5 \times b_6 \times b_7 \times b_8) - 0.371$$

where $u^*$ is the utility score for the overall health state being measured and the $b_i$’s are substituted from a table of coefficients provided by the Health Utilities Index developers for the appropriate attribute and level. For development of the single-attribute and multiplicative multiattribute utility scoring algorithms, a random sample of 504 general population adults living in the city of Hamilton, Ontario, Canada, had previously been asked to value selected health states using both a visual analog scaling technique and a standard gamble instrument. Additional details on the utility
algorithms for the Health Utilities Index Mark III are reported elsewhere.33,34

Ethics Approval
The South and West Multi-center Research Ethics Committee, England, approved this study, and participating principal caregivers provided written informed consent.

Statistical Analysis
A detailed statistical analysis plan was used to compare the health status and health-related quality of life preference-based outcomes of the children with bilateral permanent childhood hearing impairment and the children with normal hearing. The children with bilateral permanent childhood hearing impairment were stratified according to severity of hearing loss on the basis of their most recent audiologic records. The levels of severity comprised moderate (40–69 dB HL), severe (70–94 dB HL), and profound (≥95 dB HL) hearing impairment according to 4-frequency averaging of the pure-tone thresholds from 500 to 2000 Hz (or, if pure-tone thresholds were unavailable, sound fields and electrophysiologic test results).

Differences in baseline characteristics between the children with hearing impairment and the children with normal hearing were tested using the Pearson χ² test. For each of the 8 attributes of the Health Utilities Index Mark III, we compared the proportion of children with suboptimal levels of function (defined as below level 1 function) using Fisher’s exact test for equality of proportions. Differences in the single-attribute and multiattribute utility scores were tested using 2-sample t tests for unequal variance. Differences in the distributions of the multiattribute utility scores were tested using the nonparametric 2-sample Kolmogorov-Smirnov test. Finally, we performed a Tobit regression to explore the effects of severity of hearing impairment (none, moderate, severe, and profound) on the Health Utilities Index Mark III utility score. Tobit regression was required to account for the censoring of the dependent variable, the utility score, which has an upper value of 1.0. Other covariates included in the model were experience of other significant medical condition and disorders, such as cerebral palsy (no and yes), and social class of the head of household (higher occupations, intermediate occupations, lower occupations, never worked, and long-term unemployed). Because many comparisons were made, we used a conservative P value threshold of <.01 as a qualitative indicator of statistical significance. Statistical analyses were conducted by using Stata 9.0 (Stata Corp, College Station, TX).

RESULTS
A total of 168 children with bilateral permanent childhood hearing impairment and without a known postnatal cause of the impairment were identified among the 156,733 children born between 1993 and 1996 in the 4 districts in the Wessex region or between 1992 and 1997 in the 4 districts in the Greater London region. Of these 168 children, 6 were not approached because this was deemed inappropriate by their audiologist as a result of acute problems with their health at the time of recruitment, and 2 could not be traced. Of the remaining 160 children, 120 (75%; 67 boys and 53 girls) of mean age 7.9 years (range: 5.4–11.7 years) and their principal caregivers agreed to participate. Fifteen (9%) children and their principal caregivers refused study participation, and 25 (16%) children and their principal caregivers did not respond to repeated study requests. No statistically significant differences were found between participants and nonparticipants with respect to age at follow-up, gender, and severity of hearing loss. A group of 63 (19.4%) of the 325 approached children who had normal hearing (37 boys and 26 girls) and a mean age of 8.1 years (range: 6.3–9.8 years) and their principal caregivers also agreed to participate.

No statistically significant differences were identified in the distributions of age at follow-up, gender, mother’s highest educational qualification, social class of the head of household, or use of English as the first language at home between the children with bilateral permanent childhood hearing impairment and the children with normal hearing (Table 1). Of the 120 study children with bilateral permanent childhood hearing impairment, 65 (54.2%) had moderate hearing impairment, whereas 29 (24.2%) and 26 (21.7%) had severe and profound hearing impairment, respectively. Bilateral permanent childhood hearing impairment had been confirmed by 9 completed months of age in 57 (47.5%) children, between 10 and 18 completed months of age in 27 (22.5%) children, and at a later point in 36 (30.0%) children. Eighty-six (71.6%) of the 120 study children with bilateral permanent childhood hearing impairment used oral communication, 16 (13.3%) used signing, 11 (9.2%) used a combination of oral and signing, and 7 (5.8%) used nonverbal communication and gestures. Thirteen (10.8%) of the 120 study children with bilateral permanent childhood hearing impairment had cochlear implants, 4 (3.3%), 1 (0.8%), and 8 (6.7%) of whom had moderate, severe, and profound hearing impairment, respectively. Cerebral palsy, learning disabilities, and asthma were identified in 5 (4.2%), 8 (6.7%), and 12 (10.0%) of the children with bilateral permanent childhood hearing impairment, respectively. The frequencies (proportions) of these diagnosed conditions in the children with normal hearing were 0 (0%), 0 (0%), and 5 (7.9%), respectively.

Comparisons of the frequency and proportion of suboptimal levels of function between the children with bilateral permanent childhood hearing impairment and the children with normal hearing are shown in Table 2 for each of the 8 attributes of the Health Utilities Index...
Mark III. In 6 of the 8 attributes (vision, hearing, speech, ambulation, dexterity, and cognition), there were significantly higher proportions of suboptimal levels of function among children with bilateral permanent childhood hearing impairment ($P < .01$); however, there were no statistically significant differences in the proportions of suboptimal levels of function in the emotion ($P = .539$) and pain ($P = .832$) attributes between the 2 groups of children.

When compared with the children with normal hearing, the single-attribute utility scores were significantly lower for the children with hearing impairment as a group in 6 of the 8 attributes of the Health Utilities Index Mark III (vision, hearing, speech, ambulation, dexterity, and cognition; $P < .01$; Table 3). This was also true for the moderate, severe, and profound severity subgroups for the hearing and speech attributes ($P < .01$). In addition, among the children with severe hearing impair-
ment, the single-attribute utility score was significantly lower for the cognition attribute when compared with the children with normal hearing (P < .01).

Table 4 presents descriptive statistics for the multiattribute utility scores for the comparison groups. These multiattribute utility scores summarize population preferences for the overall health state of the child across the 8 attributes of the Health Utilities Index Mark III. The mean multiattribute utility score for the children with hearing impairment as a cohort was 0.626, compared with 0.920 for the children with normal hearing, a mean difference in utility score of 0.294 that was statistically significant (P < .001). The mean multiattribute utility scores were also significantly lower in each of the 3 severity subgroups than in the group with normal hearing (P < .001).

Table 5 presents the distributions of multiattribute utility scores for the children with hearing impairment and children with normal hearing across the utility scale of the Health Utilities Index Mark III. The differences in the distributions of the multiattribute utility scores between the children with hearing impairment as a group and the children with normal hearing and between each of the severity subgroups and the children with normal hearing all were statistically significant (P < .001).

Finally, the Tobit regression revealed that after controlling for other significant medical conditions and disorders and social class of the head of household, the mean multiattribute utility score for children with moderate, severe, and profound hearing impairment was 0.283, 0.477, and 0.411 less, respectively, than that for children with normal hearing (P < .0001; Table 6).

**DISCUSSION**

This study revealed that bilateral permanent childhood hearing impairment is associated with significantly increased proportions of suboptimal levels of function and significantly lower single-attribute utility scores in 6 of the 8 attributes of the Health Utilities Index Mark III: vision, hearing, speech, ambulation, dexterity, and cognition. In addition, bilateral permanent childhood hearing impairment is associated with a 0.294 reduction in the mean multiattribute utility score, a difference that is not only statistically significant but also far exceeds the 0.030 minimally important difference in utility score postulated in the literature as clinically important for evaluative purposes.35,36

A multiple Tobit regression that was conducted as part of the analytical strategy revealed that even after controlling for the presence of other significant medical conditions and disorders, children with bilateral permanent childhood hearing impairment had lower multiattribute utility scores, on average, than children with normal hearing. Although some evidence exists to suggest that the impact of significantly reduced auditory input on adverse health outcomes might be mediated through increased risk for isolation, social phobias, and impulsivity,37 additional research is required to elucidate...
the complex biopsychosocial pathways that are likely to be involved. A surprising aspect of our results is that the decrement (reduction) in the mean multiattribute utility score associated with severe hearing impairment is higher than that associated with profound hearing impairment (0.477 vs 0.411). This finding reflects differences in resource use and costs and multidimensional health profile measures between the severity subgroups; consequently, we were unable to make any firm assertions about the probable bounds of the effects of bilateral permanent childhood hearing impairment on the utility scores. A number of studies have applied multiple regression methods to general population data from the United States to estimate the marginal disutility of chronic conditions in adults. The marginal disutility of the vast majority of chronic conditions was found to be smaller than that for the hearing impairment states revealed by our multiple Tobit regression, suggesting that our results are likely to be of relevance to clinical and policy decision makers.

There are a number of strengths to this study. First, it is based on a cohort of children drawn from defined geographic areas rather than clinic-based populations; consequently, selection biases are unlikely to represent a major problem. Second, children were recruited from 8 districts of southern England that broadly reflect the nation’s socioeconomic and underlying health characteristics; therefore, the study is likely to have high external validity. Third, the study used validated and reliable approaches to measure the health status, health-related quality of life preference-based outcomes, and severity of hearing impairment of the study children. Fourth, the analysis used a geographic- and age-based comparison group that was specifically recruited for this study rather than data from siblings, which are prone to biases as a result of continuously changing development profiles throughout childhood, or comparisons with British population norms for which limited data are available.

There are 3 broad caveats to the study findings, which should be borne in mind by readers. First, despite a rigorous process designed to recruit an appropriately sized comparison group, only 19.4% of the children with normal hearing and their principal caregivers who were approached agreed to participate in the study. However, in no case did a chi-square analysis of the background sociodemographic variables in Table 1 show a significant difference between the children with hearing impairment and children with normal hearing. Consequently, we are confident that the low response rate did not systematically bias our results. Second, each child’s health status and health-related quality of life were assessed by the

| TABLE 5 | Distribution of Health Utilities Index Mark III Multiattribute Utility Scores |
|-----------------------------|-----------------------------|-----------------------------|
| Utility Score | Severity of Hearing Impairment, n (%) | Children With Hearing Impairment (n = 120), n (%) | Children With Normal Hearing (n = 63), n (%) |
|-----------------------------|-----------------------------|-----------------------------|
| Moderate (n = 65) | Severe (n = 29) | Profound (n = 26)* |
| <0.000 | 5 (7.7) | 3 (10.3) | 1 (4.0) | 9 (7.6) | 0 (0.0) |
| 0.000–0.250 | 1 (1.5) | 3 (10.3) | 2 (8.0) | 6 (5.0) | 0 (0.0) |
| 0.251–0.500 | 6 (9.2) | 8 (27.6) | 4 (16.0) | 18 (15.1) | 2 (3.2) |
| 0.501–0.750 | 13 (20.0) | 8 (27.6) | 7 (28.0) | 28 (23.5) | 4 (6.3) |
| 0.751–1.000 | 40 (61.5) | 7 (24.1) | 11 (44.0) | 58 (48.7) | 57 (90.5) |
| a One child had a missing multiattribute utility score. |

| TABLE 6 | Health Utilities Index Mark III Multiattribute Utility Score in Relation to Severity of Hearing Impairment (Tobit Regression) |
|-----------------------------|-----------------------------|-----------------------------|
| Parameter | Adjusted Mean Difference (95% CI) | P |
| Severity of PCHI | | |
| None* | —— | —— | —— |
| Moderate | -0.283 (-0.392 to -0.174) | <.0001 |
| Severe | -0.477 (-0.611 to -0.342) | <.0001 |
| Profound | -0.411 (-0.548 to -0.275) | <.0001 |
| Effect of other independent variables: | | |
| Other medical condition | | |
| No* | —— | —— | —— |
| Yes | -0.376 (-0.495 to -0.257) | <.0001 |
| Occupation (head of household)* | | |
| Higher | —— | —— | —— |
| Intermediate | 0.000 (-0.105 to 0.104) | .996 |
| Lower | -0.208 (-0.332 to -0.093) | <.0001 |
| Never worked | -0.099 (-0.237 to 0.040) | .161 |

Pseudo-R² = 0.597. CI indicates confidence interval.

a Reference category.

* National Statistics, Standard Occupational Classification 2000, 3-class version of higher, intermediate, and lower occupations of head of household.
principally measured utility scores; however, there is no consistent evidence to suggest that parents or other principal caregivers consistently either underreport or overreport these impairments, which suggests that there are unlikely to be systematic biases in the reporting of functional impairments. Third, although the Health Utilities Index is the most widely used of the multiattribute utility measures within the childhood context, the underlying preference weights for the Mark III health status classification system were derived from a survey of Canadian adults. Recent research suggests that our approach of indirectly estimating preference-based outcomes by attaching population-derived utility scores to Health Utilities Index Mark III health states may be a poor substitute for directly measured utility scores; however, the cognitive requirements entailed in directly estimating utility scores for health states using techniques such as the visual analog, standard gamble, and time trade-off approaches precluded a direct measurement approach among our pediatric sample. Moreover, evidence from empirical studies suggests that our approach of indirectly estimating preference-based outcomes does not result in systematic biases in group differences in utility scores.

CONCLUSIONS
This study has revealed an association between bilateral permanent childhood hearing impairment and diminished health status and health-related quality of life preference-based outcomes in midchildhood. The data generated by the study provide a basis for informing service provision and for evaluating the outcomes of preventive and treatment interventions in this area. In addition, the preference-based outcomes reported in this study can be synthesized with economic data to generate cost-effectiveness data using a composite quality-adjusted life-year outcome metric. Their use in follow-up studies in midchildhood, however, is not yet routine, and more methodologic work is required, particularly with regard to eliciting descriptions and underlying preferences for the health states experienced from the children themselves.

ACKNOWLEDGMENTS
The Hearing Outcomes Project was supported by a grant from the Wellcome Trust (reference No. 061839). Dr. Petrout is supported by a United Kingdom Medical Research Council senior nonclinical research fellowship. The National Perinatal Epidemiology Unit is funded by the Department of Health, England.

We gratefully acknowledge the help and assistance received from children, families, school staff, specialist teachers of the hearing impaired, speech and language pathologists, and audiology professionals in the local area teams.

REFERENCES
18. Brazier J, Roberts J, Deverill M. The estimation of a preference-
Association Between Iron-Deficiency Anemia and Stroke in Young Children

Jonathon L. Maguire, MD*, Gabrielle deVeber, MD, Patricia C. Parkin, MD*

*Department of Pediatrics, Divisions of *Pediatric Medicine and *Neurology, and *Pediatric Outcomes Research Team, University of Toronto Faculty of Medicine, Toronto, Ontario, Canada; *Child Health Evaluative Sciences, Hospital for Sick Children Research Institute, Toronto, Ontario, Canada

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Iron-deficiency anemia occurs with a peak prevalence of 4% to 8% in children between 1 and 3 years of age. Case reports have suggested an association between iron-deficiency anemia in healthy children and ischemic stroke. Our objective was to investigate whether iron-deficiency anemia is associated with stroke in young children.

METHODS. A case-control study was conducted of case patients who were selected from the stroke registry at the Hospital for Sick Children (Toronto, Ontario, Canada) and control subjects selected from a database of healthy children who were prospectively enrolled in an outpatient setting. Children were aged 12 to 38 months and were previously healthy with no identifiable risk factors for stroke. Age, gender, mean corpuscular volume, platelet count, and hemoglobin and ferritin levels were collected. Iron-deficiency anemia was defined as a hemoglobin level of <110 g/L, mean corpuscular volume <73 fl, and serum ferritin level <12 μg/L. Stroke was defined according to clinical and radiologic criteria.

RESULTS. Case (n = 15) and control (n = 143) subjects were similar with respect to median age and percentage of boys. Case patients had a lower median hemoglobin level and mean corpuscular volume and a higher median platelet count. Iron-deficiency anemia was significantly more common among case patients (8 [53%] of 15) than control subjects (13 [9%] of 143).

CONCLUSIONS. Previously healthy children with stroke were 10 times more likely to have iron-deficiency anemia than healthy children without stroke. Furthermore, children with iron-deficiency anemia accounted for more than half of all stroke cases in children without an underlying medical illness, which suggests that iron-deficiency anemia is a significant risk factor for stroke in otherwise healthy young children. Primary prevention and early identification of iron-deficiency anemia must remain a priority.
CHILDOOD ISCHEMIC STROKE includes both arterial ischemic stroke (AIS) and sinovenous thrombosis (SVT) and occurs at a frequency of 2 to 3 per 100,000 children each year in North America. Although childhood stroke is rare, it can lead to devastating long-term consequences and occasionally to death. A prospective cohort study of 163 children who survived ischemic stroke demonstrated that >40% had persistent moderate to severe neurologic deficits. In recent years, the case fatality rate has been estimated at 10% to 20%.

Several risk factors for childhood ischemic stroke have been consistently reported in large cohort studies of children with stroke. These include sickle cell disease, cardiac structural lesions, chronic systemic disease, cerebral arterial disease, coagulation disorders, head trauma, and subacute varicella zoster infection.

Several case reports and 1 case series of 6 children suggested an association between iron-deficiency anemia (IDA) in healthy children and ischemic stroke. One large cohort study found anemia to be present in 24% of 115 previously healthy children with AIS. The strength and magnitude of this association have not been previously reported.

IDA is a relatively common disorder in developed countries, such as Canada, the United Kingdom, and the United States, occurring with a peak prevalence of 4% to 8% in children between 12 and 36 months of age. The health implications of IDA are thought to be cognitive and motor developmental delay, which may not be improved with iron therapy.

The aim of this study was to investigate whether IDA is a risk factor for childhood ischemic stroke. We conducted a case-control study to compare the prevalence of IDA in otherwise healthy young children at the time of stroke with the prevalence of IDA in age-matched healthy children.

METHODS
The sampling frame for the study was previously healthy children between 12 and 38 months of age, because this is a group with a high prevalence of nutritional IDA. A case-control study design was used, with selection of case patients and control subjects from children who were enrolled prospectively in 2 databases.

Case patients were selected from the Canadian Pediatric Ischemic Stroke Registry. Children who presented to the Hospital for Sick Children (HSC) between January 1992 and December 2004 with AIS or SVT were considered for eligibility. AIS was defined as a focal neurologic deficit of acute onset and a computed tomography scan or MRI scan of the brain showing a lesion characteristic of a focal arterial infarct in a vascular territory consistent with the neurologic presentation. SVT was defined as headache, altered level of consciousness, seizure, or focal neurologic deficit with venography (computed tomography or MRI) showing occlusion or focal reduction of flow within cerebral venous sinuses or cerebral veins. A manual review of the hospital charts of each child who met these criteria was undertaken to assess eligibility for the study. Only children who did not have sickle cell disease, chronic systemic disease (excluding IDA), malignancy, cerebral arterial disease, meningitis, encephalitis, head trauma, recent surgery, heart disease, or indwelling catheters were included as case patients. All children who were enrolled in the stroke registry were screened for prothrombotic factors, which included activity assays for antithrombin, protein C, protein S, and the lupus anticoagulant, as well as immunologic assays for anticardiolipin antibody and molecular assays for the presence of factor V Leiden and the G20210A mutation in the prothrombin gene. Children with any of these factors identified after the stroke were not excluded; case patients with acute minor infections were also not excluded.

Control subjects were selected from a database of children who were prospectively enrolled in an observational study of iron deficiency. Children were recruited from a university-affiliated community-based pediatric practice in Toronto between November 2002 and March 2004 while attending a well-child visit. Exclusion criteria were acute febrile illness at enrollment (which may falsely increase the serum ferritin level, which was the primary outcome measure in the study), history of illness or medication associated with iron deficiency or anemia, previous diagnosis of iron depletion or anemia, current use of iron supplements, and breast milk or formula as primary source of milk at or within 8 weeks of enrollment (which provides more available iron). For the purposes of this study, children were included when blood work had been obtained.

For both case patients and control subjects, the following variables were collected: age, gender, mean corpuscular volume (MCV), platelet count, and hemoglobin and serum ferritin levels. For case patients, hematologic parameters that were obtained before the initial administration of intravenous fluids or intravenous medications either at HSC or at the child’s referring hospital were abstracted from hospital charts. For control subjects, all variables were obtained concurrently with enrollment in the observational study. IDA was defined as serum ferritin level of <12 μg/L, MCV of <73 fL, and hemoglobin level of <110 g/L. Thrombocytosis was defined as a platelet count of >450 × 10^9/L.

Baseline characteristics of the 2 groups were compared using the Pearson χ² test and the Mann-Whitney U test for gender and age, respectively. Hemoglobin level, MCV, and platelet count were compared as continuous variables by the Mann-Whitney U test. The association between stroke and IDA and thrombocytosis were calculated using odds ratios (ORs) and 95% confidence intervals (CIs). Hemoglobin levels, MCV, and platelet count were compared as continuous variables by
using the Mann-Whitney U test. Significance was defined as \( P \leq .05 \). Binomial logistic regression was used to examine interaction between variables. Data analysis was performed by using SPSS 13 (SPSS Inc, Chicago, IL). Data collection for the stroke registry and the community cohort study was approved by the HSC Research Ethics Board, and informed parental consent was obtained.

### RESULTS

Fifty-six previously healthy children who were 12 to 38 months of age presented to HSC with stroke between January 1992 and December 2004. Fifteen (27%) of these children were generally healthy before stroke and were included as case patients (Table 1). Reasons for exclusion were heart disease (\( n = 15 \)), meningitis (\( n = 5 \)), malignancy (\( n = 5 \)), head trauma (\( n = 5 \)), cerebral arterial disease (\( n = 3 \)), nephrotic syndrome (\( n = 2 \)), mastoiditis (\( n = 2 \)), and other (\( n = 4 \)). Initial hemoglobin level, MCV, and platelet count were available for all case patients. Serum ferritin level was available for all cases with hemoglobin levels of <110 g/L and MCV of <73 fL.

A total of 332 healthy children who attended a well-child visit at the community practice met full eligibility criteria, and 143 of these had blood work available and were included as control subjects. Data for platelet counts were missing for 10 control subjects.

There were no significant differences in median age (24 vs 21 months) or percentage of boys (73% vs 65%) for case patients as compared with control subjects. Mean and median hemoglobin levels were significantly lower in case patients than in control subjects (88 and 92 g/L vs 116 and 120 g/L, respectively; \( P < .01 \)). Mean and median MCVs were also significantly lower in case patients than in control subjects (65 and 57 fL vs 76 and 77 fL, respectively; \( P < .01 \)). The mean and median platelet counts were significantly higher in case patients than in control subjects (458 and 415 \( \times 10^9 \)/L vs 358 and 351 \( \times 10^9 \)/L, respectively; \( P = .05 \)). IDA was significantly more common among case patients (8 of 15 [53%]) than control subjects (13 of 143 [9%]; OR: 12; 95% CI: 4–37).

Thrombocytosis was also more common among case patients (7 of 15 [47%]) than control subjects (21 of 133 [16%]; OR: 5; 95% CI: 2–14). Using logistic regression to control for platelet count, the association between IDA and stroke remained significant (adjusted OR: 10; 95% CI: 3–33). There was no statistically significant interaction between IDA and thrombocytosis.

Three case patients with stroke were identified with previously described prothrombotic risk factors. All had subacute varicella infection, 1 of whom was also heterozygous for the factor V Leiden mutation, and none had IDA. Six case patients with stroke were noted to have acute minor infection (acute otitis media, pneumonia, acute gastroenteritis, or upper respiratory tract infection), 4 with IDA and 2 without IDA. Five case patients with stroke and IDA had SVT, 2 had AIS, and 1 had both. All 7 case patients with stroke and without IDA had AIS (Table 1).

### DISCUSSION

The results of this case-control study suggest that previously healthy children who develop vasoocclusive stroke are 10 times more likely to have IDA than healthy children who do not develop stroke. Furthermore, children with IDA accounted for more than half of stroke cases in children without an underlying medical illness, suggesting that IDA is a significant risk factor for stroke in otherwise healthy children.

Of note, SVT was more closely associated with IDA than was AIS. Although a previously reported case series of 6 children with IDA and stroke showed equal num-

---

**TABLE 1** Demographics, Stroke Type, Hematologic Parameters, and Thrombotic Risk Factors Among Case Patients

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Age, mo</th>
<th>Gender</th>
<th>Stroke Type</th>
<th>Hemoglobin, g/L</th>
<th>MCV, fL</th>
<th>Platelet Count, ( \times 10^9 )/L</th>
<th>Ferritin, ( \mu g/L )</th>
<th>IDA</th>
<th>Minor Infection</th>
<th>Thrombotic Risk Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23</td>
<td>M</td>
<td>AIS</td>
<td>113</td>
<td>85</td>
<td>394</td>
<td>—</td>
<td>—</td>
<td>AGE</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>M</td>
<td>AIS</td>
<td>79</td>
<td>57</td>
<td>656</td>
<td>5.2</td>
<td>+</td>
<td>AOM</td>
<td>—</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>M</td>
<td>AIS</td>
<td>92</td>
<td>56</td>
<td>625</td>
<td>5.1</td>
<td>+</td>
<td>Varicella and het FVLM</td>
<td>—</td>
</tr>
<tr>
<td>4</td>
<td>31</td>
<td>F</td>
<td>SVT</td>
<td>48</td>
<td>50</td>
<td>760</td>
<td>0.9</td>
<td>+</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>F</td>
<td>AIS</td>
<td>94</td>
<td>77</td>
<td>717</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>F</td>
<td>AIS</td>
<td>114</td>
<td>72</td>
<td>183</td>
<td>—</td>
<td>—</td>
<td>Pneumonia</td>
<td>—</td>
</tr>
<tr>
<td>7</td>
<td>29</td>
<td>M</td>
<td>AIS</td>
<td>60</td>
<td>51</td>
<td>622</td>
<td>2.6</td>
<td>+</td>
<td>—</td>
<td>URTI</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>M</td>
<td>AIS</td>
<td>123</td>
<td>84</td>
<td>311</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>9</td>
<td>38</td>
<td>M</td>
<td>SVT</td>
<td>76</td>
<td>50</td>
<td>415</td>
<td>3.3</td>
<td>+</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>32</td>
<td>F</td>
<td>AIS</td>
<td>120</td>
<td>78</td>
<td>481</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>11</td>
<td>24</td>
<td>M</td>
<td>AIS</td>
<td>124</td>
<td>82</td>
<td>282</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>12</td>
<td>15</td>
<td>M</td>
<td>SVT</td>
<td>64</td>
<td>51</td>
<td>493</td>
<td>11.9</td>
<td>+</td>
<td>—</td>
<td>AGE</td>
</tr>
<tr>
<td>13</td>
<td>29</td>
<td>M</td>
<td>SVT</td>
<td>29</td>
<td>47</td>
<td>333</td>
<td>11.0</td>
<td>+</td>
<td>—</td>
<td>URTI</td>
</tr>
<tr>
<td>14</td>
<td>16</td>
<td>M</td>
<td>Both</td>
<td>69</td>
<td>52</td>
<td>242</td>
<td>4.0</td>
<td>+</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>15</td>
<td>37</td>
<td>M</td>
<td>AIS</td>
<td>126</td>
<td>78</td>
<td>334</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

M indicates male; F, female; AGE, acute gastroenteritis; AOM, acute otitis media; URTI, minor upper respiratory tract infection; Varicella, varicella infection within the past 6 months; het FVLM, heterozygosity for the factor V Leiden mutation; —, data not available.
bers of children with AIS and SVT,13 our data may indicate that IDA is a greater risk factor for SVT than for AIS.

The strength and magnitude of several other risk factors for stroke have been identified. As reported in the published literature, these include activated protein C resistance (OR: 9.4; 95% CI: 2.0–44.6), elevated lipoprotein(a) (OR: 7.2; 95% CI: 3.8–13.8), antiphospholipid antibodies (OR: 6.08; 95% CI: 1.5–24.3), prothrombin gene variant (OR: 4.7; 95% CI: 1.4–14.6), factor V Leiden mutation (OR: 4.3; 95% CI: 2.8–6.5), subacute varicella infection (estimated OR: 3), and methylenetetrahydrofolate reductase (MTHFR) mutation (OR: 2.64; 95% CI: 1.5–3–33) IDA, with an OR of 10 (95% CI: 3–33) for stroke, is at least as strong a risk factor as these other risk factors. Furthermore, IDA was more common than other prothrombotic risk factors in this population of children with stroke (8 of 15 [53%] vs 3 of 15 [20%]). We were unable to assess whether IDA is an independent or additive risk factor.

Three mechanisms to explain an association between IDA and childhood ischemic stroke have been suggested: a hypercoagulable state directly related to iron deficiency and/or anemia; thrombocytosis secondary to IDA; and anemic hypoxia, whereby a mismatch between oxygen supply and end-artery oxygen demand leads to ischemia and infarction.13 Our data support an association between thrombocytosis and stroke. Healthy children who develop stroke may be 5 times more likely to have thrombocytosis than children who do not develop stroke; however, no statistically significant interaction between IDA and thrombocytosis was found. IDA seems to be an independent risk factor for stroke.

This study has several limitations. There may have been differences between case patients and control subjects other than stroke. First, children were excluded as control subjects when they had a history of breast milk or formula intake in the previous 8 weeks or an acute febrile illness but were not excluded as case patients for these reasons. Breast milk and iron-fortified formula are considered to be protective against IDA, and acute illness may falsely increase the serum ferritin level; therefore, it is more likely that these biases would lead to an underestimation of the magnitude of the association between IDA and stroke. Data on nutrition practices of case patients were not available, and we were not able to explore causal mechanisms for the development of IDA. Second, because of the rarity of childhood ischemic stroke, the time frame for enrollment of case patients was necessarily longer than that for control subjects (January 1992 through December 2004 versus November 2002 through March 2004). A change in the frequency of IDA between 1992 and 2002 could introduce bias; however, we believe that this is not likely because the published rate of iron deficiency and IDA among toddlers in Canada and the United States has not appreciably changed during this interval.14,20,27 Finally, the results from this study at most pertain to children in the age range explored, that is, 12 to 38 months, and the proportion of young children who have IDA and are at risk for stroke cannot be determined by this study.

CONCLUSIONS

Although IDA is thought to be an indolent disease with insidious long-term consequences, it may hold a 10-fold increased risk for acute stroke in well toddlers and may account for half of all strokes in otherwise well children of this age group. This research supports additional development of strategies aimed at the primary prevention and early detection of IDA in young children.

ACKNOWLEDGMENTS

This study was supported by a grant-in-aid from Danone Institute of Canada (community cohort) and grants from the Heart and Stroke Foundation of Ontario and the Bloorview Children’s Hospital Foundation (stroke registry). The Pediatric Outcomes Research Team is supported by a grant from the Hospital for Sick Children Foundation. Dr deVeber is the recipient of a Stroke Investigator Award from the Heart and Stroke Foundation of Ontario. The funding bodies had no role in design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript.

We thank Trenna Sutcliffe, MD, Sheila Jacobson, MD, Michael Peer, MD, and Shelley Westergard, RN, for recruitment of control subjects.
Impact of Computerized Prescriber Order Entry on the Incidence of Adverse Drug Events in Pediatric Inpatients

Mark T. Holdsworth, PharmD, BCOP, Richard E. Fichtl, PharmD, Dennis W. Raisch, PhD, Adrianne Hewryk, PharmD, Maryam Behra, PharmD, Elena Mendez-Rico, PharmD, Cindy L. Wong, MD, Jennifer Cohen, MD, Susan Bostwick, MD, Bruce M. Greenwald, MD

OBJECTIVES. This study was conducted to determine the impact of a computerized physician order entry system with substantial decision support on the incidence and types of adverse drug events in hospitalized children.

METHODS. A prospective methodology was used for the collection of adverse drug events and potential adverse drug events from all patients admitted to the pediatric intensive care and general pediatric units over a 6-month period. Data from a previous adverse drug event study of the same patient care units before computerized physician order entry implementation were used for comparison purposes.

RESULTS. Data for 1197 admissions before the introduction of computerized physician order entry were compared with 1210 admissions collected after computerized physician order entry implementation. After computerized physician order entry implementation, it was observed that the number of preventable adverse drug events (46 vs 26) and potential adverse drug events (94 vs 35) was reduced. Reductions in overall errors, dispensing errors, and drug-choice errors were associated with computerized physician order entry. There were reductions in significant events, as well as those events rated as serious or life threatening, after the implementation of computerized physician order entry. Some types of adverse drug events continued to persist, specifically underdosing of analgesics. There were no differences in length of stay or patient disposition between preventable adverse drug events and potential adverse drug events in either study period.

CONCLUSIONS. This study demonstrated that a computerized physician order entry system with substantive decision support was associated with a reduction in both adverse drug events and potential adverse drug events in the inpatient pediatric population. Additional system refinements will be necessary to affect remaining adverse drug events. Preventable events did not predict excess length of stay and instead may represent a sign, rather than a cause, of more complicated illness.
It has been demonstrated that adverse drug events (ADEs) contribute to morbidity and mortality in hospitalized patients. Among the few published studies conducted in hospitalized pediatric patients, these events were found to occur frequently, and many were the result of prescribing errors. It has also been shown that medication errors in hospitalized children occur at a similar rate as in adults but that the rate of potential ADEs was ~3 times greater in the pediatric population. There are few studies that have used methodology to either prevent or mitigate ADEs, particularly in the pediatric population.

Emerging data from studies of computerized prescriber order entry (CPOE) systems in adult inpatients demonstrate that computer technology can substantially reduce medication error rates, as well as the rates of potential ADEs. However, the impact of these systems on the occurrence of ADEs and patient outcomes has not been adequately investigated, particularly in hospitalized pediatric patients. Children may be more vulnerable to medication errors because of the need for a weight-based dosing approach, the increased sensitivity of children to relatively small dosing errors, alterations in pharmacokinetics and pharmacodynamics, and the lack of effective communication between children and health care personnel. Studies of CPOE in children have demonstrated a reduction in some types of medication errors, without impacting rates of ADEs.

User-developed enhancements to commercially available CPOE software applications provide for real-time feedback on the completeness and accuracy of medication orders. Prospective decision support, including dose calculations, alerting functions, and required data input, are evolving in an attempt to prevent medication errors and ADEs at every stage of the medication management process. This study was conducted to compare the incidence and types of ADEs and potential ADEs in hospitalized children before and after the implementation of a CPOE system with substantive clinical decision support.

METHODS
This study used the same prospective methodology that was used in a previous ADE study conducted on the same patient care units during 2000–2001. An ADE was defined as an injury from a medicine or lack of an intended medicine (eg, omission of an indicated medication). A potential ADE was defined as an error that had the potential to result in at least a significant injury. Potential ADEs included errors detected before drug administration, as well as errors that were administered without causing significant adverse consequences. A preventable ADE was defined as all of the ADEs that were associated with a medication error. These ADEs and potential ADEs may have been preventable by CPOE or any other medication management system or strategy. Errors that were corrected before the medication being entered onto the medication administration record, ostensibly prevented by pharmacists, physicians, and nurses, were excluded during both the pre-CPOE and post-CPOE data-collection periods. Although errors that were the result of omissions were included, errors that were the result of delays were not specifically collected.

All of the pediatric patients admitted to either the PICU or the general pediatric unit at the New York Presbyterian Hospital, Weill Cornell Medical Center, Komansky Center for Children’s Health consecutively between April 1, 2004, and October 5, 2004, were included in this study. For comparison purposes, data from a previous published study on these same units during the pre-CPOE period from September 2000 to May 2001 were also included. The PICU and the general pediatric unit are located within a large metropolitan tertiary care center and are composed of 20 and 30 beds, respectively. The PICU is a multidisciplinary unit that functions as a major referral center for a large health care network and a major cancer center and provides medical and surgical intensive care to all of the pediatric patients, including infants.

A single clinical pharmacist, serving as the primary reviewer during each study period, prospectively identified events and potential events. The primary reviewer identified recordable events on a daily basis in a prospective manner via review of physician and nursing notes, pharmacy records, medication administration records, and laboratory data. In addition, nursing, medical, and pharmacy personnel were interviewed to resolve questions raised during medical chart review. Nursing and pharmacy supervisory personnel were also interviewed weekly to obtain any reports of additional adverse events. This was the sole job function of the reviewer during the time of data collection. Demographic and other case-specific data, including case-mix index (CMI; an index used to assess case severity), were collected for all of the admitted patients. These data served as the denominator for all of the further analysis.

The proximal cause and systems failure were assigned by the primary reviewer using previously published definitions. Examples of proximal causes include lack of drug knowledge, lack of information about the patient, and drug stocking. Examples of systems failures include drug-knowledge dissemination, dose and frequency standardization, and medication order tracking. A rating of event severity was assigned by the primary reviewer according to a previously published scale as significant, serious, or life threatening. This severity rating was primarily determined by the magnitude of dosing error, therapeutic index of the drug, and route of administration. Determination as to whether an ADE was the result of a medication error, categorization of error
types, and the validation process for events were identical to a process used previously. This process involved the independent assessment by 2 secondary reviewers and consensus among the primary and secondary reviewers regarding all of the events. Events that were not the result of a medication error (adverse drug reactions) were also included for comparison purposes.

Data from the previous ADE study conducted during 2000–2001 (pre-CPOE) were used as the baseline data for comparison purposes. The only substantive global change on these 2 patient care units between the time of completion of the previous study and initiation of the current investigation was the introduction of a CPOE system with decision support. There were limited enhancements of the unit dose drug distribution system (eg, enhanced unit dose preparation for some oral liquids) and other minor improvements to the medication management system (eg, override use enhancement, addressing look-alike sound-alike issues, and ongoing formulary revision). These enhancements were consistent with evolutionary changes generated by any organization actively focusing on ongoing quality improvement. However, these changes were minor compared with the revolutionary change experienced with the implementation of the CPOE system.

The CPOE module for medication order entry within the electronic medical chart (Eclipsys System 2000, SCC 1.3 MO1; Eclipsys, Boca Raton, FL) was modified from the commercial product. A list of ~200 medications commonly used in pediatric patients was selected for incorporation into the system. A collaborative effort among pharmacists, physicians, informatics specialists, nurses, and performance improvement specialists was undertaken to create a pediatric dosing table (PDT) that was used as a source database for the dosing recommendations triggered by the system after drug selection. Data reflected in the PDT were adapted from standard pediatric and neonatal references and expert opinion. Each drug in the PDT had ≥1 line to stratify recommendations based on combinations of parameters, such as dosage form, gestational age, postnatal age, and weight and/or body surface area. The PDT was not stratified for physiologic function or disease state.

Default doses were developed to address the most common indication for any particular drug. For drugs with multiple indications and/or variable dosing recommendations, the system provided users with alternative dosing information in a text box incorporated into the order default screen. In addition, the system populated all of the medication order fields in the order default screen with the recommended dosing parameters outlined in the PDT. Doses that exceeded a preestablished maximum triggered an alert to the prescriber but did not alter the order or prevent the order from being executed. A dose-rounding function served to round doses to a degree appropriate for pediatric and neonatal patients. An alert was triggered notifying the prescriber when the rounded dose varied >5% from the originally calculated dose. The dosing weight for each patient was a required field within the system, and the system did not permit medication order entry without this required information. The system also used standardized concentrations for medicated intravenous infusions. Pediatric generalists, specialists, and intensivists reviewed and approved the content of the PDT, and the process and content were approved by the institution’s formulary and therapeutics committee and medical board. The CPOE system was developed, approved, extensively tested, and implemented over a period of 20 months after the preliminary (pre-CPOE) data-collection period. The majority of this time was spent on gaining consensus related to appropriate defaults within the PDT. System refinements and user acclimation occurred for an additional 15 months before the follow-up (post-CPOE) data-collection period. This development process led to the post-CPOE study period commencing ~36 months after the pre-CPOE period.

Orders from the CPOE module were electronically interfaced to the pharmacy system (Cerner Pharmnet; Cerner Corporation, Kansas City, MO) where the order was reviewed by a pharmacist and transmitted to the medication profile in the automated dispensing machines (Pyxis Medication Station 2000; Cardinal Health, Dublin, OH) located on patient care areas. In addition, active orders were directly transmitted to the medication administration record within the electronic medical chart. The study was approved by the organization’s institutional review board.

Statistical Analysis

Each patient admission was categorized as having a preventable ADE, potential ADE, or no ADE. The number of preventable and potential ADEs per 100 admissions and per 1000 patient-days were determined for pre-CPOE and post-CPOE time periods. These values and other categorical data were compared between study periods using χ² analyses and, when appropriate, determination of relative risk (RR) with 95% confidence intervals (CIs). These RR values were used to determine the number needed to treat estimates. Specifically, this was calculated as the inverse of the absolute differences in rates of potential and preventable ADEs per admission and per patient-day. The number needed to treat 95% CIs were calculated by the same procedure. Comparisons of continuous data between the 2 study periods were performed using analysis of variance (across all 3 of the ADE types) or Student’s t tests between time periods within each ADE type. Variables that were significantly different between the 2 study periods within the preventable and potential ADE groups were entered into a logistic regression model. This allowed for identification of the impact of CPOE implementation after controlling for
other variables. RRs and 95% CIs were calculated for each variable, with the no-ADE group as the comparator category. In each analysis, an α level of .05 was used to test for statistical significance.

RESULTS
Data for the post-CPOE period were collected from 1210 patient admissions between April and October 2004. Data for the pre-CPOE period were previously published and were collected from 1197 consecutive patient admissions between September 2000 and May 2001. Data for these 2 populations are provided in Table 1. Small differences were noted in the age, number of medications, CMI, and inpatient distribution between the patient care units for the no-ADE population. During the post-CPOE period, there was an increased proportion of preventable ADEs and a decreased proportion of potential ADEs in the ICU. The length of stay (LOS), CMI, and distribution of patients within each nursing unit were also different between the 2 time periods for the potential ADE group.

Preventable ADE and potential ADE incidence rates are also provided in Table 1. The RRs of an adverse event are further depicted in Fig 1, by both event type and severity. There was a significant reduction in total ADEs, preventable ADEs, and potential ADEs after the implementation of the CPOE system. The RR of a preventable ADE was 0.56 (95% CI: 0.34–0.91), and that of a potential ADE was 0.37 (95% CI: 0.25–0.55) after the implementation of CPOE. Furthermore, there were reductions in all of the significant events (combination of preventable ADEs and potential ADEs) between the pre-CPOE (n = 127) and post-CPOE time periods (n = 44; RR: 0.34; 95% CI: 0.24–0.49). Similar reductions were found for all of those events rated as serious or life-threatening between the 2 time periods (pre-CPOE [n = 13] and post-CPOE [n = 3]; RR: 0.23; 95% CI: 0.07–0.80).

Using a number needed to treat analysis, it was determined that CPOE was associated with the avoidance of 1 potential ADE every 20.2 admissions (95% CI: 13.9–30.0 admissions) and 1 preventable ADE every 59.0 admissions (95% CI: 36.2–96.1 admissions). Similarly, a potential ADE was avoided every 212 patient-days (95% CI: 144–313 patient-days), and a preventable ADE was avoided every 865 patient-days (95% CI: 534–1400 patient-days). Using the average daily census observed during the post-CPOE study period (41.5 patients), CPOE was associated with the avoidance of a potential ADE every 5 days and a preventable ADE every 20 days.

Logistic regression analyses were used to determine the impact of CPOE on rates of potential and preventable ADEs, after controlling for LOS, number of medications, CMI, and patient care unit (Table 2). To perform the analysis, the number of medications and CMI were converted into categorical variables by dividing them into quartiles relevant to each period (cutoff points for quartiles shown in Table 2). CPOE remained a significant independent predictor of decreased occurrence of preventable ADEs (odds ratio [OR]: 0.76; 95% CI: 0.59–0.97) and potential ADEs (OR: 0.56; 95% CI: 0.46–0.70). The results also showed that, in addition to CPOE, exposure to the top quartiles of number of medications (pre-CPOE: >8 medications; post-CPOE: >10 medications) and CMI (pre-CPOE: >2.85; post-CPOE: >2.43) were significant predictors of potential and preventable adverse events.
The most common error types are presented in Table 3. The character of the errors changed after the implementation of the CPOE system, with a notable reduction in overall errors, dispensing errors, and drug-choice errors. However, some types of dosing errors, particularly underdoses, continued to represent a substantial proportion of the total errors. The continued prevalence of underdose errors was particularly evident in the ADE group. The most common type of preventable event in both time periods was inadequate analgesia (pre-CPOE period: \( n = 28 \); post-CPOE: \( n = 29 \)). All of the other types of preventable events were a minority (ie, \( \leq 5 \) events in both data sets). After analgesics, the drug class most associated with preventable events was antibiotics. After CPOE implementation, there were reductions in adverse events associated with certain antibiotic drug classes, including aminoglycosides (12 vs 0) and cephalosporins (14 vs 2). The most common types of adverse events that were decreased in relation to aminoglycosides were potential ADEs (\( n = 11 \)) related to medication order tracking (\( n = 6 \)) and dose and frequency standardization (\( n = 5 \)). For cephalosporins, the reductions in adverse events again involved potential ADEs (\( n = 13 \)) most commonly related to medication order tracking (\( n = 8 \)). Problems with medication order tracking before CPOE were because of the pharmacy not receiving dis-

**TABLE 2** Logistic Regression Results for Occurrence of Preventable or Potential ADEs With No ADE as a Reference

<table>
<thead>
<tr>
<th>Variable</th>
<th>Wald Statistic</th>
<th>Preventable ADEs (( n = 69 ), Adjusted OR (95% CI))</th>
<th>Potential ADEs (( n = 125 ), Adjusted OR (95% CI))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient care unit (reference: PICU)</td>
<td>4.99 (0.082)</td>
<td>0.76 (0.58–0.99)</td>
<td>0.89 (0.72–1.10)</td>
</tr>
<tr>
<td>CPOE system (reference: pre-CPOE)</td>
<td>31.24 (&lt;0.001)</td>
<td>0.76 (0.59–0.97)</td>
<td>0.56 (0.46–0.70)</td>
</tr>
<tr>
<td>Medications (reference: lowest quartile)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top 25% (pre-CPOE &gt; 8, post-CPOE &gt; 10)</td>
<td>39.03 (&lt;0.001)</td>
<td>1.63 (1.01–2.64)</td>
<td>2.78 (2.00–3.87)</td>
</tr>
<tr>
<td>26%-50% (pre-CPOE = 6–8; post-CPOE = 7–9)</td>
<td>1.14 (0.573)</td>
<td>1.25 (0.76–2.01)</td>
<td>1.13 (0.77–1.67)</td>
</tr>
<tr>
<td>51%-75% (pre-CPOE = 3–5; post-CPOE = 4–6)</td>
<td>9.77 (0.008)</td>
<td>1.52 (0.96–2.41)</td>
<td>0.59 (0.39–0.87)</td>
</tr>
<tr>
<td>CMI (reference: lowest quartile)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top 25% (pre-CPOE &gt; 2.85, post-CPOE &gt; 2.43)</td>
<td>11.77 (0.003)</td>
<td>1.73 (1.10–2.73)</td>
<td>1.59 (1.12–2.25)</td>
</tr>
<tr>
<td>26%-50% (pre-CPOE = 1.17–2.85; post-CPOE = 1.06–2.43)</td>
<td>1.08 (0.483)</td>
<td>1.26 (0.81–1.96)</td>
<td>0.97 (0.67–1.40)</td>
</tr>
<tr>
<td>51%-75% (pre-CPOE = 0.72–1.17; post-CPOE = 0.71–1.06)</td>
<td>1.55 (0.460)</td>
<td>0.86 (0.50–1.48)</td>
<td>0.78 (0.51–1.20)</td>
</tr>
<tr>
<td>LOS (continuous)</td>
<td>22.14 (&lt;0.001)</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Adjusted ORs indicate likelihood of classification as preventable ADE or potential ADE in the comparator versus the reference level of each variable. NA indicates not applicable.
continuation orders and subsequently continuing to disperse discontinued medication.

There were some differences in the proximal cause and systems failures between the 2 time periods. Events resulting from lack of drug knowledge were less prevalent in the post-CPOE group (n = 34) compared with the pre-CPOE group (n = 54; RR: 0.62; 95% CI: 0.40–0.96). Drug stocking as a cause of these events was eliminated after CPOE implementation (35 vs 0; P < .001 [χ²]).

The most frequent systems failures attributed to these events pre-CPOE were dose and frequency standardization (25.4%), medication order tracking (24.6%), and dose and identity checking (9.7%). After CPOE implementation, errors associated with dose and scheduling were reduced, and there was a shift in systems failure attribution for these events to drug-knowledge dissemination (60%) and medication order tracking (10%).

There were no differences in patient disposition among patients with preventable ADEs or potential ADEs between pre-CPOE and post-CPOE time periods. Patients with preventable ADEs or potential ADEs were less likely to be routinely discharged and more likely to be discharged to either another institution or to care under a home health care program in the pre-CPOE time period. There was also a difference in patient disposition between the no-ADE groups in the 2 time periods, with the post-CPOE group having a greater chance for non-routine discharge. However, further analysis of the no-ADE group between the 2 time periods did not reveal significant differences for individual types of discharge disposition (eg, discharge to another institution, discharge to home health care, or deceased).

**DISCUSSION**

The results of this study demonstrate that implementation of a CPOE system with substantive decision support on inpatient pediatric patient units was associated with a reduction in both preventable ADEs and potential ADEs. The reduction in these adverse events was most likely related to the implementation of the CPOE system within the institution, because this was the only major process modification between the 2 time periods. Although there were some minor differences between the 2 data sets for some of the study variables (eg, CMI, LOS, and patient care unit distribution), logistic regression findings demonstrate that, after controlling for these variables, CPOE remained a significant predictor of lower adverse event rates.

The LOS remained nearly identical between the ADE and potential ADE groups within each time period, but the ~7-day reduction in the post-CPOE period does suggest that patients with adverse events or potential adverse events in the latter study group may have been different. A previous study at this institution in this same population determined that exposure to >8 medications was a strong predictor of these events. Consequently, this variable was compared in the no-ADE groups between the 2 time periods to explain the difference in LOS. It was shown that the percentage of patients receiving >8 medications increased from 16.1% to 31.9% (P < .001) in the no-ADE group after CPOE implementation. This indicated that the change in adverse event incidence between the 2 time periods was associated with a shift of patients at higher risk for an event to the no-ADE group rather than an actual difference in disease severity between the pre-CPOE and post-CPOE populations. In addition, LOS in the overall patient cohort was not different. One possible explanation for this finding is that those patients with more complicated illnesses were less likely to experience an adverse event after CPOE implementation and that such patients were being discharged more promptly.

In addition to an overall decrease in adverse events, specific types of errors decreased after implementation of the CPOE system, including drug-choice errors and dispensing errors. These changes are likely a direct result of the CPOE system. For instance, the majority of dispensing errors within the potential ADE category were because of a lack of diligence in communicating discontinuation orders, particularly related to discontinued antibiotics, in the manual system used previously. With the advent of the CPOE system, the discontinuation order became automated, and this potential system failure was addressed.

Underdosing of analgesic medications continued to represent a large number of the total dosing errors after the implementation of CPOE. This dosing error was also noted in the pre-CPOE data set with approximately the same incidence. The opiate analgesic doses were coded in the PDT using the lower end of the standard dose range from pediatric references with the intention for users to titrate doses based on pain scale assessments. The system was not designed to improve the manual system and to use rules recommending dose escalations for single or sequential suboptimal pain scores. This may have contributed to the continued observation of inadequate analgesia in the post-CPOE population. The development of continuous pain score assessment and dosing escalation recommendations through alerting may
be a worthwhile endeavor for implementation and future study. Underdosing of systemic analgesics is a well-documented problem in the pediatric arena and will likely require decision support and cultural change to produce a significant improvement in analgesic dosing.\textsuperscript{20,21}

This study also confirms findings from a previous study in pediatric inpatients.\textsuperscript{6} Preventable ADEs did not necessarily predict excess resource use. When compared with patients with potential ADEs, it was demonstrated that the mean LOS was not significantly greater for the preventable ADE group during either time period. In addition, LOS was substantially greater for both of these groups compared with the no-ADE group. Because patients with potential ADEs have not actually experienced an injury from a drug, these data suggest that excess resource use, as reflected by LOS in this group, likely reflects a more complicated patient. This is also supported by greater CMI values and medication exposures in the ADE and potential ADE populations. It is apparent that both of these groups are more complicated than most pediatric inpatients and that these differences do not seem to be because of ADEs themselves. ADEs may be a result of complicated illness rather than a causal factor in the pediatric inpatient population.

Several studies of CPOE have been performed in recent years in both the adult and pediatric populations.\textsuperscript{8,9,12,13,22–25} Most CPOE studies have noted a reduction in either medication errors and/or potential ADEs, but none have been reported to reduce the overall rate of ADEs or preventable ADEs.\textsuperscript{8,9,11–13,22–25} However, I previous study has demonstrated that a computer-assisted antibiotic management program led to reductions in errors and adverse events.\textsuperscript{26}

The majority of these studies have not included clinical outcomes such as ADEs as measured variables but rather surrogate outcomes, such as errors, laboratory test use, and compliance.\textsuperscript{11} The lack of a documented impact on ADEs in the majority of these studies is thought to be because of the fact that these are rare events and represent a distinct minority (~1%) of all medication errors. For instance, a previous time series analysis of the impact of CPOE on medication errors and events found only 5 ADEs during the baseline period and a total of 19 such events during 3 different follow-up periods that were each 7 to 10 weeks in length.\textsuperscript{8} Furthermore, many studies may not have been adequately powered to detect a difference in the rates of these relatively rare events. The fact that this study has demonstrated a difference in preventable ADEs and potential ADEs with CPOE likely reflects the number of events captured through rigorous prospective identification rather than relying on the low yield from a voluntary reporting system. Negative results of the impact of CPOE on ADEs have been reported when ADE incidence is based on voluntary reporting, likely because of the limited number of events identified with a voluntary reporting system.\textsuperscript{12}

Implementation of CPOE has also not routinely demonstrated success in improving patient outcomes. In particular, a few recent studies have actually demonstrated a continued high incidence of ADEs, increased mortality, or facilitation of medication errors after implementation of CPOE.\textsuperscript{22,23,25} The reasons for CPOE failure were postulated to be multifactorial, although common themes within these studies were hastened implementation timelines, lack of clinical decision support, and lack of adequate process design. Assessing outcomes after a prolonged system refinement and acclimation phase in addition to the development of extensive decision support were likely key factors that contributed to the findings of this study. The most important decision support modifications were likely the provision of pertinent patient demographic information during the ordering process, drug knowledge in order-detail screens, dosing recommendations, and notification to the pharmacy of discontinuation orders. Similar findings may not be reproducible with commercially available CPOE systems with nominal decision support. It would seem logical that the less comprehensive the dosing decision support the less likely a CPOE system will have the ability to reduce preventable ADEs. The risk of failing to customize existing systems to assist with prescribing for pediatric patients is likely substantial. The complexity of decision support development, however, may preclude its development in institutions where substantive resources are not available to dedicate to these functions. Highly advanced systems will require substantial resource commitment on an ongoing basis from a wide variety of providers (eg, pharmacists, physicians, information technology personnel, and nurses) focusing on system monitoring, assessment, expansion, and improvement on an ongoing basis.

Using new technologic advances to reduce the potential occurrence of ADEs is perhaps more important in the pediatric population than for any other patient group. Children are particularly vulnerable to errors within the health care system.\textsuperscript{5,12} Although consequences of this vulnerability are not well studied, emerging evidence indicates that the most severe outcomes may be amplified in the pediatric population. An analysis of mortalities associated with medication errors demonstrated that children aged 0 to 9 years represented the second greatest number and percentage of deaths (after patients aged 70–79 years) and the largest number of life-years lost when distributed by 10-year age cohorts.\textsuperscript{27,28} Thus, minimizing errors in the pediatric population may be of heightened importance, because the pediatric population’s vulnerability to dosing errors is significant and the downstream benefit in terms of life-years gained from successful error prevention is greater than for the adult population. Despite these critical benefits from ADE pre-
vention in children, it is unfortunate that all of the commercially available CPOE systems require significant modification to effectively serve pediatric populations.

This research did have some limitations. Because we used an observational pre-post comparison, differences in outcomes between the 2 study periods could represent a difference in patient populations or other factors, including potential bias in the post-CPOE data-collection period, that were not accounted for in the logistic regression. Possible confounding variables could have been the minor medication management enhancements discussed previously. Another limitation of this study was the potential for interreviewer variability between the 2 study periods, because the sole primary reviewers were different individuals in the 2 study periods. This potential variability was mitigated by having the reviewer for the first study period train and review cases with the primary reviewer for the second study period. In addition, the data-collection periods were not seasonally matched, because the pre-CPOE data collection was performed from September through May. However, based on the current ADE literature, there is no evidence that adverse event incidence is subject to seasonal variation. Because the study did not collect data on errors that were corrected before medication being entered into the medication administration record, it is unknown whether certain types of these errors may have also been impacted by the CPOE system. These errors were excluded because the purpose of the study was to examine the incidence of ADEs and potential ADEs that were not being prevented by the existing patient care system during both data-collection periods. In addition, errors that were the result of delays were not a specific focus of this research and would require additional inquiry to determine the impact of CPOE on delays in therapy.

CONCLUSIONS
This study demonstrated that a CPOE system with substantive decision support was associated with a reduction in both ADEs and potential ADEs among pediatric inpatients. Some ADEs continued to occur in this patient population, particularly underdosing of analgesic medications, suggesting that additional system modifications will be necessary to affect the remainder of these events. Preventable events did not predict excess resource use and may have instead represented a sign, rather than a cause, of more complicated illness.

REFERENCES
25. Han YY, Caricillo JA, Venkataraman ST, et al. Unexpected increased mortality after implementation of a commercially
MOST SCIENCE STUDIES APPEAR TO BE TAINTED BY SLOPPY ANALYSIS

“We all make mistakes and, if you believe medical scholar John Ioannidis, scientists make more than their fair share. By his calculations, most published research findings are wrong. Dr. Ioannidis is an epidemiologist who studies research methods at the University of Ioannina School of Medicine in Greece and Tufts University in Medford, MA. In a series of influential analytical reports, he has documented how, in thousands of peer-reviewed research papers published every year, there may be so much less than meets the eye. These flawed findings, for the most part, stem not from fraud or formal misconduct, but from more mundane misbehavior: miscalculation, poor study design or self-serving data analysis. ‘There is an increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims,’ Dr. Ioannidis said. ‘A new claim about a research finding is more likely to be false than true.’ The hotter the field of research the more likely its published findings should be viewed skeptically, he determined. Take the discovery that the risk of disease may vary between men and women, depending on their genes. Studies have prominently reported such sex differences for hypertension, schizophrenia and multiple sclerosis, as well as lung cancer and heart attacks. In research published last month in the *Journal of the American Medical Association* Dr. Ioannidis and his colleagues analyzed 432 published research claims concerning gender and genes. Upon closer scrutiny, almost none of them held up. Only one was replicated. Statistically speaking, science suffers from an excess of significance. Overeager researchers often tinker too much with the statistical variables of their analysis to coax any meaningful insight from their data sets. ‘People are messing around with the data to find anything that seems significant, to show they have found something that is new and unusual,’ Dr. Ioannidis said. In the US, research is a $55-billion-a-year enterprise that stakes its credibility on the reliability of evidence and the work of Dr. Ioannidis strikes a raw nerve. In fact, his 2005 essay ‘Why Most Published Research Findings Are False’ remains the most downloaded technical paper that the journal *PloS Medicine* has ever published.”


Noted by JFL
Ontogeny of Bilirubin-Binding Capacity and the Effect of Clinical Status in Premature Infants Born at Less Than 1300 Grams

George Jesse Bender, MD, William James Cashore, MD, William Oh, MD

Department of Pediatrics, Brown Medical School, Women and Infants’ Hospital of Rhode Island, Providence, Rhode Island

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

BACKGROUND. Bilirubin is toxic to the brain and enters the brain in unbound form. Serum unconjugated, unbound bilirubin may be a good predictor of bilirubin encephalopathy. Unbound bilirubin levels may depend on the bilirubin-binding capacity of albumin, which has not been described for neonates of <28 weeks’ gestation.

OBJECTIVE. The purpose of this work was to determine the ontogeny of bilirubin-binding capacity and the effect of clinical status in very preterm neonates.

METHODS. A total of 152 neonates (23–31 weeks’ gestational age; 440–1300 g) were enrolled prospectively. At 5 days of age, total serum bilirubin and unbound bilirubin were measured with the unbound bilirubin-A1 analyzer (Arrows Co, Osaka, Japan) and albumin with the Bromocresol-purple method. Scatchard plots were used to estimate bilirubin-binding affinity and capacity. Clinical status for each infant was rated as high, moderate, or low risk by using a modified Score for Neonatal Acute Physiology model. Low risk was considered clinically stable.

RESULTS. Unbound bilirubin has a significant, direct correlation to total bilirubin and is greater in unstable than in stable neonates. For the entire cohort, bilirubin-binding capacity had a direct relationship to gestational age. The bilirubin-binding capacities of infants in the low- and high-risk groups also had a direct relationship to gestational age. Bilirubin-binding capacity was greater in the low-risk group (20.8 ± 4.6 mg/dL; 356 ± 79 μmol/L) than in the moderate- (17.8 ± 3.5 mg/dL; 304 ± 60 μmol/L) or high- (17.3 ± 3.4 mg/dL; 296 ± 58 μmol/L) risk groups. Bilirubin-binding affinity did not differ by clinical risk status or gestational age.

CONCLUSIONS. In very preterm, very low birth weight infants, bilirubin-binding capacity is directly proportional to gestational age. Bilirubin-binding capacity is lower and unbound bilirubin higher in unstable than in stable neonates. These data may be useful in guiding the management of hyperbilirubinemia in very low birth weight infants.
Although extreme hyperbilirubinemia has been associated with acute bilirubin encephalopathy and kernicterus, the long-term impact of modest hyperbilirubinemia in very low birth weight neonates is unknown. Treatment protocols for phototherapy and exchange transfusion use total serum bilirubin (TSB) as the parameter. Wennberg et al recently concluded that experimental and clinical data suggest that measurement of unbound bilirubin (UB) in newborns with hyperbilirubinemia may improve the risk assessment for neurotoxicity. Cellular uptake of UB varies in different tissues. UB may vary widely for a given TSB, because the binding capacity of albumin may be altered by many clinical and physiologic variables. Bilirubin-binding capacity (BBC) has been described for sick and well neonates of >28 weeks’ gestation but not for the very preterm and extremely preterm neonates. With increasing survival of these infants, particularly with the retrospective association between peak serum bilirubin and neurodevelopmental impairment, there is a need to clarify the risk of hyperbilirubinemia in this group.

The primary hypothesis of this study was that BBC is directly related to gestational age and is influenced by clinical status in a very preterm and very low birth weight population. The secondary hypothesis was that BBC is greater (and, hence, UB levels lower at a given TSB) in clinically stable than in unstable neonates at any gestational and postnatal age.

METHODS

Subjects

Neonates born at or transferred to Women and Infants Hospital of Rhode Island, with birth weight between 401 and 1300 g and estimated gestational age from 23 weeks and 1 day through 30 weeks and 6 days, were eligible and enrolled when parental consent was obtained. Gestational age was determined by the best obstetrical estimate of first-trimester fetal ultrasound for fetal dating, last menstrual period, or Ballard scores. Infants were excluded if they had lethal chromosomal, craniofacial, or other congenital abnormalities or if they died within 48 hours of life. The study underwent Women and Infants’ Hospital Institutional Review Board review and approval, and informed consent was obtained.

Between February 2003 and November 2004, 152 neonates were enrolled. One pair of twins was subsequently withdrawn per parental request. Blood samples were collected, spun, and stored according to protocol with sufficient serum for serial binding capacity measurements from 137 neonates. Twenty-four samples did not exhibit 2-site binding, having linear low-affinity binding even at low bilirubin levels, hence making BBC estimation unreliable. Eight charts did not have enough data for risk assessment, leaving 105 risk-classed patient sera with BBC estimates.

Study Design

Only 1 serum sample was collected per infant to minimize the physiologic impact of blood withdrawal. This was collected on day 5 ± 1 (SD) of life, at the expected peak of TSB, taken from routine blood draws during the evaluation of hyperbilirubinemia. Each sample was 0.3 to 0.5 mL of whole blood, shielded from light, and centrifuged, and then the serum was separated and frozen at −20°C within 2 hours of extraction for later analysis. Samples were labeled with a unique identifier for storage, with the code correlating to patient identification kept in a protected logbook.

Identity was blinded during total bilirubin, UB, and albumin measurement and analysis. Physiologic data were collected, and clinical risk status was assigned, without knowledge of UB measurement or BBC calculations. Care decisions for the infant were made according to standard management with indirect/total bilirubin and without knowledge of UB, BBC, or assigned clinical risk status.

Samples were thawed within 1 hour of UB determination. UB and TSB were measured concurrently using the UB-A1 analyzer (Arrows Co, Osaka, Japan) at the recommended sample dilution (1:43), at 29°C, with buffer and UB and TSB standards from the Arrows Reagent kit. Measurement error was estimated by repeat testing in triplicate when enough serum was available; SE was 0.086 mg/dL for TSB and 0.017 μg/dL for UB. Concurrent serum albumin concentration was measured with the Bromocresol-purple method as described by Pinnell and Northam. UB was measured at increasing TSB concentrations, with 2 μL of solution (0.001 g of bilirubin-9a from Frontier Scientific [Logan, UT] suspended in 25 μL of 0.1 M EDTA and 10 μL of 10 N NaOH, then diluted with 1 mL of deionized H2O, set >1 hour, split, and measured at 3 concentrations) per 25 μL of patient sera. For sera with clearly distinct high-affinity binding affinities represented by the Scatchard plot (5–7 points), an estimate of primary site binding affinity (slope of K1) was done, and binding capacity (K1 intercept of the x-axis) was calculated from the raw data. Thus, BBC was an estimate of the total bilirubin at which saturation of the primary binding site on albumin would occur in the absence of secondary and nonspecific binding sites.

The presence of multiple binding sites was determined by Scatchard transformation of UB versus incremental total bilirubin:albumin molar ratios. Typically, a very high affinity primary binding (K1) was noted, followed by lower affinity sites (K2 and K3) at increasing total bilirubin levels. For these samples, BBC and affinity calculations were made. A subset of these sera (n = 15) had sufficient data points for confirmation of the K1 estimate with nonlinear regression, using GraphPad Prism (GraphPad, San Diego, CA). For some sera with high TSB at baseline (n = 14), K1 estimation was not...
possible, because primary binding sites were saturated. Other sera (n = 10) exhibited low binding affinity even at low total bilirubin levels. This “K₂ analysis,” segregating the sera into those demonstrating primary versus secondary site binding at the onset, was done independently by 2 investigators, using assessment of the Scatchard plot linearity, initial total bilirubin:albumin molar ratio, deviation from population mean affinities (K₁ vs K₂), and deflection point between dominant site regression lines.

Clinical Data Collection
Perinatal and neonatal clinical data included gestational age derived by best obstetrical estimate, Apgar scores at 5 minutes, gender, race, and birth weight. Physiologic variables for risk assessment during the first and second 24-hour periods after admission to the NICU were recorded, including (1) axillary temperature, (2) lowest mean arterial blood pressure by transducer connected to an arterial catheter or Dynamap, (3) blood gas data with mean arterial blood pressure by transducer connected to a blood gas analyzer, and (4) average urine output in milliliters per kilogram per hour as measured by nursing staff during routine care. The presence of seizures observed by a caretaker in the first 24 hours was also noted. Small-for-gestational-age status was considered as birth weight less than the 3rd percentile for gestational age.

Infants were classified by clinical risk, stratified into a categorical, ordinal variable with 3 possible values: high, moderate, or low risk. The risk class was determined by clinical and physiologic data, augmented by the concurrent degree of intervention provided by the clinicians. Measures of hemodynamics, oxygenation, temperature, metabolism, acid/base, and renal status were recorded and then processed through a weighted risk algorithm designed to balance the importance of each variable. The weighting is a simplification from the validated Score for Neonatal Acute Physiology Perinatal Extension (SNAP-PE)-II logistic model for mortality risk in the perinatal period. The same 5 physiologic variables and cutoff parameters were used in the current risk assignment as in the SNAP-PE model. On an a priori basis, a score from 0 to 5 was assigned for each variable, in proportion to the weights in the original multiple regression model. Additional points were added based on birth weight, small-for-gestational-age status, low 5-minute Apgar score suggesting difficult transition, or persistent seizures in the first 24 hours life consistent with hypoxic-ischemic encephalopathy. The weighted risk scoring is summarized in Table 1. The net risk score was the sum of all of the risk points, which was labeled “high risk” for a score of ≥8, “moderate risk” for a score of 4 to 7, and “low risk” for a score of ≤3. These risk classifications correlate with a 10%, 1% to 10%, and <1% mortality risk in the SNAP-PE study, respectively.

Sample Size
Cashore et al³ found that the difference in BBC between sick versus well was larger in the term than in preterm neonates. The projected effect size in this study, 2.1 mg/dL, was divided among 3 risk groups, with assumed uneven distribution (high risk: 15.5; moderate risk: 16.2; low risk: 17.6), with SDs at 0.6 and 0.5 from sick and well low birth weight neonates, respectively. A priori grouping of infants into 3 gestational age categories was done at 23 to 25, 26 to 27, and 28 to 30 weeks. Three gestational age clusters for 3 risk groups (9 cells), an α value of .05, and a power of 90% result in a net sample size of 135 patients. As noted above, the actual number of risk-adjusted BBC data points is 105, which precluded the intended cluster analysis but was sufficient for logistic regression.

Data Analysis
Risk-class grouping differed according to the time interval of assessment. Among the 105 neonates with sufficient serum and risk-assessment data, 30%, 36%, and 34% were at low, moderate, and high risk, respectively.

### Table 1: Weighted Risk Score

<table>
<thead>
<tr>
<th>Score</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low temperature, °F</td>
<td>&gt;96</td>
<td>95–96</td>
<td>20–29 or p</td>
<td>&lt;95</td>
<td>&lt;20 or P</td>
<td>—</td>
</tr>
<tr>
<td>Mean blood pressure, mm Hg</td>
<td>&gt;30</td>
<td>10–25</td>
<td>0.3–0.9 or H</td>
<td>&lt;7.10</td>
<td>—</td>
<td>&lt;0.3</td>
</tr>
<tr>
<td>Po₂/FiO₂ ratio</td>
<td>&gt;2.5</td>
<td>1.0–2.5</td>
<td>0.3–0.9 or H</td>
<td>&lt;7.10</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Lowest serum, pH</td>
<td>&gt;7.25</td>
<td>7.10–7.25</td>
<td>—</td>
<td>&lt;7.10</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Urine output, mL/kg per h</td>
<td>&gt;1.0</td>
<td>0.1–1.0</td>
<td>—</td>
<td>&lt;0.1</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Multiple seizures in first 24 h</td>
<td>No</td>
<td>—</td>
<td>—</td>
<td>Yes</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Birth weight, g</td>
<td>&gt;1000</td>
<td>—</td>
<td>750–1000</td>
<td>&lt;750</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Weight &lt;3rd percentile</td>
<td>AGA</td>
<td>—</td>
<td>SGA</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>S-min Apgar score</td>
<td>≥7</td>
<td>—</td>
<td>—</td>
<td>&lt;7</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

* Indicates minimal (single-agent ≤5 µg/kg per hour) pressor support or volume expansion (≤10 mL/kg); P, greater pressor support or volume expansion; H, high-frequency ventilation; fraction of inspired oxygen; AGA, appropriate for gestational age; SGA, small for gestational age; —, no data.

* Data indicate the most extreme physiologic derangement in the first 12 hours of NICU admission.

* Data indicate the average urine output over the first 12 hours of NICU admission.
when classified by using the first 24 hours of physiologic data. Results were the same whether the first 24-hour data or average 48-hour data were used: the tables and figures, therefore, present only the first 24-hour data.

Linear regression analyses (Pearson) were performed to compare UB and BBC among the risk groups, with means compared by analysis of variance and Bonferroni multiple comparisons. Statistical analysis was done using Graphpad Prism. Risk-group analysis within gestational age clusters was not significant, in part because of inadequate numbers of low-risk/low-gestation neonates.

RESULTS

Demographics and Clinical Data

As shown in Table 2, the neonates at high risk are of significantly lower gestational age and birth weight and more likely to be boys. There is no difference in the distribution of the Hispanic and black population among the risk groups.

Outcomes that proxy for severity of illness separate into groups as expected. Neonates at high risk are more likely to receive surfactant, be ventilated at 48 hours, and require oxygen at 36 weeks’ after conception age. The raw SNAP-PE scores differ significantly between the risk groups. No patient-specific outcomes are measured.

Total and Unbound Serum Bilirubin

The study infants’ actual serum total and UB data, before titration, are summarized in Table 3. Total bilirubin is not different between the groups, but UB shows a significant (analysis of variance: \( P < .01 \)) increase with increasing risk. Albumin is statistically greater (analysis of variance: \( P < .001 \)) in the low-risk group than in the other 2 risk groups but relatively equivalent from a clinical standpoint. Gestational age has no significant effect on albumin, total bilirubin, or UB in any of the risk groups (data not shown). As shown in Fig 1, the direct correlation between UB and TSB is significant \(( r^2 = 0.376; \ P < .001 \), as well as within each of the risk groups: low \(( r^2 = 0.230; \ P < .001 \), moderate \(( r^2 = 0.389; \ P < .001 \) and high \(( r^2 = 0.486; \ P < .001 \). The linear regression between unbound and total bilirubin has a different slope in each risk group (F test, \( P = .03 \)), and each slope is significantly nonzero (F test; \( P < .001 \). The SEs of replicate measurements were 0.003 ± 0.002 g/dL of albumin, 0.21 ± 0.14 mg/dL of TSB, and 0.04 ± 0.04 μg/dL of UB.

Bilirubin-Binding Capacity

Figure 2 shows the relationship between gestational age and BBC expressed as milligrams per deciliter of bilirubin with different risk groups. As a group, irrespective of risk category, BBC has a direct relationship to gestational age \(( r^2 = 0.187; \ P < .001 \). This relationship to gestational age holds in the low-risk \(( r^2 = 0.179; \ P < .05 \) and high-risk \(( r^2 = 0.139; \ P < .05 \) groups but not the moderate-risk group. The slopes of the regression lines are not significantly different. BBC is greater (analysis of variance, \( P < .001 \)) in the low-risk group \((20.8 ± 4.6 \text{ mg/dL}; 356 ± 79 \text{ μmol/L})\) than in the moderate- \((17.8 ± 3.5 \text{ mg/dL}; 304 ± 60 \text{ μmol/L})\) or high- \((17.3 ± 3.4 \text{ mg/dL}; 296 ± 58 \text{ μmol/L})\) risk groups.

Bilirubin-Binding Affinity

Binding affinity (mean \( K_1 \) slope) is not significantly different by analysis of variance between clinical risk groups and gestational age (data not shown). The 3 \( K \) values are distinct from each other \((K_1; 11.4 ± 5.9 [95\% confidence interval: 10.3–12.4]; K_2; 2.6 ± 1.4 [95\% confidence interval: 2.3–2.8]; and \ K_3; 1.2 ± 0.88 [95\% confidence interval: 0.9–1.4]; all values are \times 10^7 \text{ L·mol}^{-1}).

DISCUSSION

Although the incidence of kernicterus has decreased significantly with the widespread use of phototherapy and exchange transfusions over recent decades, an increase has been noted particularly in the near-term and term infants over the past few years.10 Guidelines for appropriate bilirubin levels at which to initiate such interventions have been updated periodically for near-term or term infants.11 Lack of good scientific evidence for very preterm neonates resulted in publication of guidelines projected from term TSB data,13 although kernicterus has been reported in premature neonates at low TSB levels.15 Comparative neuropathy study has shown that the localized patterns in term kernicterus may differ from the more diffuse uptake in premature neonates with hyperbilirubinemia.16 Similar patterns have been shown by MRI for term17 and preterm infants with kernicterus.18 Peak serum bilirubin has been retrospectively correlated with hearing impairment and Bayley Psychomotor Development Index of <70.4

Our data support the relationship between UB and
TSB that has been historically assumed in the clinical management of very low birth weight infants with mild hyperbilirubinemia. Neonates at high risk have a greater UB at a given TSB than neonates at either low or moderate risk, which justifies the use of clinical status to tailor interventions. However, the difference in mean UB between the risk groups is statistically but not clinically impressive, which, along with a high UB-TSB coefficient of variation, suggest that measurement of UB itself may be of even more use. Using the same methodology in a general population of newborns with jaundice, a UB between 0.86 and 1.19 µg/dL correlated with an increased likelihood of kernicterus,\(^19\) consistent with previous studies showing reversible acute bilirubin encephalopathy at UB levels of >1.0 µg/dL,\(^20\) as well as increased risk of kernicterus.\(^{21,22}\) Although our population mean UB is in the nontoxic range, 31%, 16%, and 7%, respectively, of the patients at high, moderate, and low risk had a UB level of >1.0 µg/dL.

For unclear reasons, a subset (14 with high initial TSB and 10 with low TSB) of our patients have no identifiable \(K_i\) binding, because only a single shallow slope was observed on Scatchard plot. Both the mean TSB (10.2 mg/dL) and UB (0.89 µg/dL) are higher for these patients

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Actual Serum Bilirubin, UB, and Albumin Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Risk Strata</td>
</tr>
<tr>
<td>No.</td>
<td>Low</td>
</tr>
<tr>
<td>TSB, mean ± SD, mg/dL</td>
<td>8.85 ± 2.64</td>
</tr>
<tr>
<td>TSB, mean ± SD, µmol/L</td>
<td>151 ± 45</td>
</tr>
<tr>
<td>UB, mean ± SD, µg/dL</td>
<td>0.53 ± 0.28</td>
</tr>
<tr>
<td>UB, mean ± SD, nmol/L</td>
<td>9.1 ± 4.8</td>
</tr>
<tr>
<td>Albumin, mean ± SD, g/dL</td>
<td>2.75 ± 0.29</td>
</tr>
<tr>
<td>Albumin, mean ± SD, µmol/L</td>
<td>404 ± 42</td>
</tr>
</tbody>
</table>

NS indicates not significant; —, no data.

FIGURE 1
The relationship between UB and TSB was significant overall and in each of the risk groups (low, moderate, and high).

FIGURE 2
BBC has a direct relationship to gestational age. This relationship to gestational age holds in the low-risk and high-risk groups but not in the moderate-risk group, when assessed at 24 hours. The slopes of the regression lines are not significantly different.
than for the other patients. Although binding affinity and capacity measurements could, therefore, not be made for these sera, these neonates are potentially at the greatest risk for bilirubin encephalopathy. Fully 37% of these sera with no BBC measurement have an initial UB level of >1.0 µg/dL.

The majority of patients in the current study have BBC estimates higher than those retrospectively associated with kernicterus using the Sephadex method. Clinical or postmortem evidences of kernicterus were not observed in our patients. BBC is directly proportional to gestational age, which is not accounted for by significant relationships of albumin, TSB, or UB with gestational age. As expected, UB has a much closer relationship with TSB than either gestational age or BBC. Second, the significance of BBC versus gestational age in the high- and low-risk groups is lost in the moderate-risk group, which may represent a gray zone in the risk stratification. Third, because >80% of the variance in BBC is not related to gestational age, other serum components or clinical factors must have an important role. Competitors for albumin binding include many drugs and molecules, such as fatty acids. Our patients received exogenous free fatty acids via parenteral nutrition. However, we did not measure serum fatty acid concentrations. Analysis of BBC versus raw risk score does not identify which clinical factors contribute most (data not shown).

Bilirubin-binding affinity is highly variable within our population and does not correlate with clinical risk severity or gestational age. The mean affinity is >1 order of magnitude higher than that reported for term newborns with other methods. The UB-A1 analyzer uses a kinetic peroxidase method that may underestimate the actual serum UB concentration. UB measurement can be complicated by dilution of competitors, interference by conjugated bilirubin, and failure to correct for the rate-limiting albumin-bilirubin dissociation step. Interference with UB measurement by bilirubin photoisomers must also be considered given the prevalence of phototherapy in this population, although the effect in the clinical range is expected to be minimal. A combined peroxidase-diazoo assay adjusting for these factors found UB to be poorly correlated with, and consistently greater than, that measured with the kinetic peroxidase method used in this study. A fivefold underestimation of the mean UB using the standard peroxidase method has been shown by using minimal dilution methods. Thus, the next generation of UB measurement may be more precise to bring UB to the forefront of hyperbilirubinemia management.

CONCLUSIONS

Within the gestational age range of 23 to 31 weeks and birth weight between 401 and 1300 g, the total BBC of the primary binding sites of albumin is directly proportional to maturity. BBC is greater in stable than in unstable neonates. UB is lower in stable than in unstable neonates. UB correlates strongly with total bilirubin and is affected by clinical status. No consistent relationship was seen between binding affinity and gestational age or physiologic risk. These data may be useful in guiding the management of hyperbilirubinemia in these infants, particularly if the National Institute of Child Health and Human Development Neonatal Network Phototherapy Trial identifies significant relationships between developmental outcome and either UB or BBC.

ACKNOWLEDGMENT

Dr Oh was supported by grant 2 U10 HD27904 from the National Institute of Child Health and Human Development Neonatal Research Network.

REFERENCES

8. Graphpad Prism 4 for Windows [computer program]. San Diego, CA: GraphPad Software, Inc; 2005
In the past, the real advances in medicine have often come not from research specifically directed against a target but rather from discoveries made in fields other than the one being studied. In cancer, chemotherapy arose from the development of instruments of chemical warfare, the study of nutritional disease, and the effect of electric current on bacterial growth. A major tumor-suppressor gene was discovered through research on polio vaccines. Stem cells were discovered through research on bone-marrow transplants for leukemia and for whole-body irradiation from nuclear weapons. Agents for alteration of mood and behavior came from attempts to combat surgical shock, the treatment of tuberculosis, and the search for antibiotics. When scientists were allowed to pursue whatever they found, serendipitous discovery flourished. Today, targeted research is pretty much all there is. Yet, as Richard Feynman put it in his typical rough-hewn but insightful manner, giving more money ‘just increases the number of guys following the comet head.’ Money doesn’t foster new ideas, ideas that drive science; it only fosters applications of old ideas, most often enabling improvements but not discoveries.”


Noted by JFL
Epidemiology of Oronasal Hemorrhage in the First 2 Years of Life: Implications for Child Protection

Neil McIntosh, DSc(Med)a, Jacqueline Y. Q. Mok, MB, ChB, MDb, Adrian Margerison, MB, MA, BChirc

aDepartment of Child Life and Health, University of Edinburgh, Edinburgh, United Kingdom; bDepartment of Community Child Health, Royal Hospital for Sick Children, Edinburgh, United Kingdom; cDepartment of Community Child Health, Borders General Hospital, Melrose, Roxburghshire, United Kingdom

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

BACKGROUND. Epistaxis in childhood is common but unusual in the first years of life. Oronasal blood has been proposed as a marker of child abuse.

METHODS. We performed a retrospective review of all hospital notes of children in the Lothian region of Scotland who were <2 years of age and in whom facial blood had been recorded over a 10-year period.

RESULTS. There were 77,173 accident and emergency department attendances with 58,059 admissions during the 10-year study period in children <2 years of age; 16 cases of nose bleed and 3 cases of hemoptysis were recorded. All cases of hemoptysis were associated with significant bouts of coughing and respiratory infections. Epistaxis in 8 cases was associated with visible trauma and in 4 cases with thrombocytopenia (secondary to malignancy in 3). In 2 cases, an associated apparent life-threatening event was described, and in 2 cases there was a coincident upper respiratory tract infection. Review of previous and subsequent history suggested 7 cases of “accidental” injury that might have been caused by abuse. These cases are described here. All children who presented with this problem to the accident and emergency department had been admitted for observation or management.

CONCLUSIONS. Epistaxis is rare in the accident and emergency department and hospital in the first 2 years of life and is often associated with injury or serious illness. The investigation of all cases should involve a pediatrician with expertise in child protection.

www.pediatrics.org/cgi/doi/10.1542/peds.2007-2097
doi:10.1542/peds.2007-2097

Key Words
oronasal hemorrhage, epistaxis, pulmonary hemorrhage, hemoptysis, child abuse, infancy

Abbreviations
AED—accident and emergency department
ICD-10—International Classification of Diseases, 10th Revision
ENT—ear, nose, and throat
ALTE—apparent life-threatening event

Accepted for publication: Aug 10, 2007
Address correspondence to Neil McIntosh, DSc(Med), Department of Child Life and Health, University of Edinburgh, 20 Sylvan Place, Edinburgh EH9 1UW, United Kingdom.
E-mail: neil.mcintosh@ed.ac.uk

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
The finding of blood on the face of a young infant must be alarming for a parent. The differential diagnosis contains serious and life-threatening conditions and has also been associated with child abuse. Little has been published about the epidemiology of nose bleeds or bleeding from the mouth in early life, although it is regarded as rare. This study reports the epidemiology of oronasal bleeding in those presented to the accident and emergency department (AED) and/or were admitted at a large district general hospital for children that also provides specialty services for southeast Scotland.

METHODS

All AED attendances and admissions at the Royal Hospital for Sick Children have had their symptoms and diagnoses coded for the last 12 years (according to the International Classification of Diseases, 10th Revision [ICD-10]). Medical charts within a 10-year period (April 1, 1996, to March 31, 2006) were reviewed to find diagnostic codes that the hospital coding clerks said that they could have used to record the presence of facial blood. These codes were R04 (epistaxis/nasal bleeding/nose bleeds), R04.2 (hemoptyisis), R04.8 (pulmonary hemorrhage, older than newborn), and K13.7 (other and unspecified lesions of oral mucosa).

The notes of all patients <2 years of age and with residence in Lothian, Scotland, were reviewed independently by 3 pediatricians, 2 of whom were lead pediatricians in child protection of many years standing. Data were extracted about the incident of the hemorrhage (if confirmed) and the reasons for previous and/or subsequent admissions, and the date of the last hospital contact was noted. A cause for the bleeding was allocated (usually the cause recorded on the index admission). For cases in which there was delay in consulting medical services after an acute injury or a witness at the time of presentation considered the story to be incompatible or variable (phrases used concurrently in the notes are in quote marks in Table 1), abuse was considered possible. In addition, each of us independently made a decision as to whether abuse was possible (footnote “a” in Table 1) on the basis of both the event and the previous and subsequent history. If only 1 or no pediatrician felt that abuse was possible, abuse was said to be unlikely; if 2 or 3 pediatricians independently were concerned about possible abuse, abuse was said to be possible.

In each case we made a decision with the information available at the time on whether (by current standards) a multiagency discussion would have been appropriate during the index admission.

RESULTS

The total population for the Lothian area of Scotland is 790 000, of whom 17 000 are <2 years of age. During the 10-year review period, there were 350 000 AED attendances and 207 289 admissions from the Lothian region of southeast Scotland: 38 215 AED attendances and 34 590 admissions were of children <1 year old, and 38 958 AED attendances and 23 469 admissions were of children between the ages of 1 and 2 years. It is unlikely that infants with this problem in the Lothian region would be seen outside the Royal Hospital for Sick Children.

There were 888 attendances or admissions during this time for the R04 ICD-10 code, which includes epistaxis, pulmonary hemorrhage, and hemoptysis. There were 62 attendances or admissions for lesions in the mouth (ICD-10 code K13.7). Of the children and infants who were <2 years old, there were only 30 cases, and all were admitted to the hospital. The AED and inpatient notes of all these children were evaluated. Nine cases referred from outside the region were excluded (because we had no denominator information), which left 21 cases from within the Lothian region. Three cases of hemoptysis occurred during the second year of life (at ages 378, 499, and 502 days of life); in each case the hemoptysis occurred in an infant who had been to their general practitioner in the recent past with chronic and paroxysmal coughing. These cases were accepted as described, and there seemed to be nothing, even in retrospect, that would suggest further problems. No case of pulmonary hemorrhage outside the neonatal period occurred in a Lothian infant. Two Lothian cases coded as K13.7 were found, both in patients with malignancy who were receiving chemotherapy. The lesions were described in the notes as either mucositis or ulcers, and in neither case was there associated bleeding.

There were 16 cases of epistaxis: 8 in children <1 year old and 8 in children between 1 and 2 years old. The detail of each case is shown in Table 1. Eleven of 16 cases had a cause allocated at the time (2 secondary to an upper respiratory tract infection), and for 4 children with recurrent epistaxis a cause was never established (cases 3, 6, 9, and 16) even after ear, nose, and throat (ENT) referral. In 8 cases there was obvious associated injury. All injuries were accepted as accidental at the time of their presentation, although in 4 cases (1, 3, 7, and 14) child abuse was considered by at least 1 staff member. Review of the notes for delay in presentation, or subsequent history of poorly explained trauma, suggested that 5 injuries might have been associated with physical abuse (cases 3, 7, 8, 14, and 15). Two other children without obvious injury at presentation may also have been abused (cases 1 and 16), and 1 (case 13) was thought after review by the 2 child protection experts to be a possible case of fabricated or induced illness. Four infants had thrombocytopenia at the time of the nose bleed (1 with idiopathic thrombocytopenic purpura and 3 with malignancy and receiving chemotherapy, 1 of which was associated with injury). The child in case 1 had possibly been suffocated (overlaid) and certainly
## Cases of Epistaxis

<table>
<thead>
<tr>
<th>Case ID</th>
<th>Age, d</th>
<th>Notes on Index Incident</th>
<th>Visible Injury on Index Incident</th>
<th>Coagulation (Index Incident)</th>
<th>Associated ALTE</th>
<th>Associated URTI</th>
<th>MAD*</th>
<th>Previous/Subsequent Contact</th>
<th>Accident**</th>
<th>Possible Abuse***</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58</td>
<td>Parents sleeping with twins; father apneic; mother woke as 1 twin was crying, other twin had blood on face; matted, and was not breathing well; no history of overlaying; no delay; sepsis (?) as patchy CR on admission; possible “overlaying/possible suffocation though no history”</td>
<td>No external evidence of injuries</td>
<td>Normal</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Aged 2 y scald on dorsum of right hand from pulling hot microwave; aged 2 1/2 y fell onto coffee table, resulting in nose bleed</td>
<td>Possible</td>
<td>++ +</td>
</tr>
<tr>
<td>2</td>
<td>94</td>
<td>Fell on face from 3- to 4-ft high chair to carpet; bleeding both nostrils; no delay; nil subsequent</td>
<td>Abrasion/bruise over philtrum</td>
<td>Unknown</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No further RHSC contact; last seen for a squint at nearby hospital (ICD-10 10.1.07)</td>
<td>Probable</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>122</td>
<td>Age 5 wk; unexplained rectal bleeding, recurrent episodes from 6-wk of age; no cause found after ENT referral; “spontaneous” bruised nose when first seen at age 4 mo for epistaxis</td>
<td>Bruise on right cheek</td>
<td>Normal</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Several presentations with constipation; fell of blood in stool; fell of blood in vomit; vomiting coffee grounds; (?) bruising under eyes from time to time (ICD-10 11.1.05)</td>
<td>No</td>
<td>++ +</td>
</tr>
<tr>
<td>4</td>
<td>123</td>
<td>Presented with platelet count of 14000 and disseminated malignancies in skin, liver, and brain; final diagnosis was choriocarcinoma</td>
<td>None described</td>
<td>Low platelet count</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Regular treatment in oncology department until death</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>145</td>
<td>Fell 3 ft off kitchen worktop when mother was surprised by father; no loss of consciousness; no delay; nil subsequent</td>
<td>Bruise on forehead and nose</td>
<td>Unknown</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Several presentations with asthma/eczema (ICD-10 6.11.05)</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>255</td>
<td>Recurrent spontaneous epistaxes; cause not identified by ENF department</td>
<td>None described</td>
<td>Unknown</td>
<td>No</td>
<td>No</td>
<td>?</td>
<td>No further RHSC contact (ICD-10 12.6.98)</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>264</td>
<td>Mother on toilet; infant “fell out of bed”, dad picked him up and infant bled from nose; 2-h delay; many social problems; ALTE at 14 mo of age</td>
<td>Swollen, bruised nose, faint bruise on right lower back</td>
<td>Unknown</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Nose bleed from falling out of high chair 2 wk previously (ICD-10 12.2.07)</td>
<td>Possible</td>
<td>++ +</td>
</tr>
<tr>
<td>8</td>
<td>358</td>
<td>Holding on to coffee table and fell sideways, hitting head on chair, bled from both nostrils; no delay; similar injury (fell against cabinet) 6 mo later</td>
<td>Blood around nose</td>
<td>Unknown</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Several presentations with childhood injuries/illnesses (ICD-10 6.1.05)</td>
<td>Possible</td>
<td>c</td>
</tr>
<tr>
<td>9</td>
<td>398</td>
<td>Recurrent epistaxis from 11 mo of age; 1 started with a cold, re-referred by general practitioner; bled on bedsheet daily and nose bleeds daily during day for 4 mo; no diagnosis made by ENF</td>
<td>None described</td>
<td>Unknown</td>
<td>No</td>
<td>Yes</td>
<td>?</td>
<td>Several presentations with childhood injuries/illnesses (ICD-10 4.8.03)</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>445</td>
<td>Brush and petechial rash 1 d; platelet count of 7000, clear case of ITP</td>
<td>None described</td>
<td>Low platelet count</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Regular hematology follow up; no other problems</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>534</td>
<td>Hemophagocytic lymphohistiocytosis; platelet count of 35 000 at development of epistaxis</td>
<td>None described</td>
<td>Low platelet count</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Regular oncology follow up; no other problems</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>638</td>
<td>AML diagnosed at 3 wk; fell off chair at 20 mo; platelet count of 39000 at time</td>
<td>Swelling and bruising of nose</td>
<td>Low platelet count</td>
<td>No</td>
<td>No</td>
<td>?</td>
<td>Regular hematology follow up; no other problems or incidents</td>
<td>Probable</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>684</td>
<td>Recurrent epistaxes from 18 mo; bilateral, spontaneous, no trauma, no underlying diagnosis made</td>
<td>Unknown</td>
<td>Unknown</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Maternal anxiety; recurrent nonspecific illnesses; “unwell for years” (ICD-10 27.9.05)</td>
<td>No</td>
<td>++ - (??)</td>
</tr>
<tr>
<td>14</td>
<td>684</td>
<td>Fell down concrete stairs at 23 mo of age (not witnessed); previous signs and vomiting; 2 reflux at 4/12, “Mother-vague history”, at 8 mo “mother not coping”</td>
<td>Swollen bloody nose</td>
<td>Unknown</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Intractable constipation, noncompliant with treatment; social problems; has supervision requirement (ICD-10 6.2.07)</td>
<td>Possible</td>
<td>++ +</td>
</tr>
<tr>
<td>15</td>
<td>715</td>
<td>Fell out of buggy at 4 mo of age HI with much facial damage</td>
<td>Facial abrasions, laceration upper lip, torn frenulum</td>
<td>Unknown</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>11 mo; HI fell down concrete steps; at 17 mo: HI with nose bleed; tripped at 22 mo; nose bleeds overnight twice; recurrent attendances for nose bleeds; still 1 yr (ICD-10 30.6.05)</td>
<td>Possible</td>
<td>++</td>
</tr>
<tr>
<td>16</td>
<td>730</td>
<td>Nose bleed at 8 mo of age; no cause; at 12 mo, recurrent nose bleed; referred to ENF; no definite cause but none was cautioned</td>
<td>None described</td>
<td>Normal</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>From 1 to 5 y: 7 admissions with HI and/or nose bleeds (ICD-10 26.5.04)</td>
<td>No</td>
<td>+</td>
</tr>
</tbody>
</table>

a Diagnosis of possible abuse was made by each author independently from review of all notes.
b Decision about multiagency discussion was warranted at the time of event from the information available (? indicates that there was insufficient detail available to decide).
c Diagnosis of accident was made by consensus from review of all notes.

**+, ++, and +++ indicate the number of authors who were suspicious about abuse; URTI, upper respiratory tract infection; MAD, multiagency discussion; CR, chest radiograph; HI, head injury; RHSC, Royal Hospital for Sick Children; ITP, idiopathic thrombocytopenic purpura; AML, acute monoblastic leukemia.

### Notes

- Table 1: Cases of Epistaxis
- Possible abuse was assessed by each author independently.
- Multiagency discussion was considered based on available information.
- Possible causes were noted for each case.

### Additional Information

- Diagnosis of possible abuse was made by each author independently.
- Decision about multiagency discussion was warranted at the time of event.
- Consensus was reached for diagnosis of accidental cases.

### Definitions

- ALTE: Unexplained sudden infant death syndrome.
- URTI: Upper respiratory tract infection.
- MAD: Multiagency discussion.
had features that suggested that there had been an ALTE. In case 14, the child had a disturbed conscious level, which may have been a result of head injury or an ALTE.

The child protection experts decided, on the basis of the information available at the time of presentation of the index incident, that by current standards a multiagency discussion would have been appropriate in 50% of the presentations.

**DISCUSSION**

For 12 years all children who attended the Royal Hospital for Sick Children AED and/or were admitted have had their diagnoses coded by using the ICD-10 when they left the hospital. ICD-10 code R04 was most helpful and included pulmonary hemorrhage outside the neonatal period, epistaxis, and hemoptysis. ICD-10 code K13.7 indicated ulcers and mucositis in the mouth that, after review of the notes, were not associated with bleeding. Hemoptysis and pulmonary hemorrhage seemed to be extremely rare and genuine events.

Whereas overall 1 in 400 children who attend the AED will present with an epistaxis, attendance for this problem is rare for children under the age of 2 years (1 in 6400 <1-year-olds and 1 in 6500 1- to 2-year-olds), and it is very likely that they will be admitted. In an average year in Lothian, there are ~1.6 nosebleeds for a population of 2-year-olds (17,000), which gives 0.94 nosebleeds per 10,000 of this population.

This study shows that blood on the face from the nose or mouth in the first 2 years of life is rare and that spontaneous epistaxis is also rare. This study also confirms that this sign is associated with some serious underlying diagnoses. In the absence of physical injury, it would be appropriate to check the patient’s platelet count and perform a clotting screen as serious illness (malignancy or chronic liver disease) comes into the differential diagnosis.

Oronasal hemorrhage has been described in a number of case reports of imposed suffocation in which the perpetrator later confessed. It was part of the history of 11 of 38 parents of children with recurrent ALTEs who were seen by covert video surveillance to suffocate their infants. It proved to be a highly significant discriminator between this group and a control group found to have a natural cause for repeated ALTEs. This series, however, was highly selected, and a population denominator was impossible to define. Krous et al found facial blood present at autopsy in 7% of 406 cases of sudden and unexpected deaths in infancy, and in 14 of 28 cases the blood was present before or in the absence of cardiopulmonary resuscitation. The authors suggested that when this sign is found at autopsy, it is reasonable to question whether the diagnosis is truly sudden infant death syndrome.

The mechanism of oronasal hemorrhage is often unclear. Blunt trauma as part of accidental injury may be obvious and may also be apparent when there are other signs of physical abuse (eg, bleeding gums, pinch marks on the nose). When found in association with an ALTE, hemorrhagic pulmonary edema secondary to upper airway obstruction has to be considered. Case reports of confessed suffocation and strangulation suggest that this is a possible mechanism and usually produces frothy or pink blood-stained fluid and sometimes produces respiratory distress. Pathologists clearly recognize that the presence of either large amounts of blood in the lungs or intraalveolar siderophages in cases of sudden infant death syndrome may be a marker for imposed suffocation. For cases in which “blood and mucus” is described, it is less clear that the origin may be the lungs.

No children described in this report suffered injury from definite abuse or from injury that was sufficiently worrying to trigger child protection proceedings at the time of injury, although by current standards a multiagency discussion would have been appropriate in half of the cases. The circumstances of the index event or hindsight available from subsequent events might raise child protection anxieties in 44% of them. In the first case, it was thought more likely that it was an overlying event rather than suffocatory abuse. In many of the other cases, the injuries were in a child at the toddling age, when injuries from falls are not infrequent. It is the repeated nature, varying stories, or other oddity that suggests that reconsideration may be warranted. The pediatrician has a dilemma. Should the concerns be voiced, leading to anger and affront in the parents, or should injuries be “accepted” as accidental and inevitable in the small child, leaving them at further risk? The law in the United Kingdom is unclear, although the lay press often takes a simplistic view of what is always a complex situation.

Do general practitioners refer such cases to ENT departments rather than pediatricians? All children who are referred for ENT consultation in Lothian are evaluated at the Royal Hospital for Sick Children and are entered into our database. Only 2 cases (6 and 9) of the 16 were seen in the ENT department; in neither case was a diagnosis made.

Two other issues need to be addressed to determine the completeness of these discussions and the final impact of the problem, neither of which can be answered by this study. First, do parents of a young child discovered with facial blood or a nose bleed always seek medical help? If there is a strong family history of nose bleeds in older children, will the parents manage it from their previous experience and a knowledge that ensuing harm is unlikely? Second, do primary care practitioners always refer a child of this age with facial blood for a hospital opinion? We have no data to answer either of these questions, but it is noteworthy that all infants who attended our AED during this period and with this prob-
lem were admitted for observation, which indicates that the AED staff took each individual event seriously.

CONCLUSIONS
Facial blood from a nose or mouth bleed or pulmonary hemorrhage is rare in the first 2 years of life. The problem should be taken seriously by medical staff, because there are important underlying conditions that may warrant urgent evaluation and management. A pediatrician with some experience in child protection should always be involved, and cases should not be referred directly to ENT surgeons.

ACKNOWLEDGMENTS
We thank Tony Wright, for the 4D data analysis, and the Coding and Records staff at the Royal Hospital for Sick Children (Edinburgh).

REFERENCES
Effects of Corticosteroid on Henoch-Schönlein Purpura: A Systematic Review

Pamela F. Weiss, MDa,b,c, James A. Feinstein, MDb, Xianqun Luan, MSd, Jon M. Burnham, MD, MSCEa,b,c, Chris Feudtner, MD, PhD, MPHb,c,e

aDivision of Rheumatology, bPediatric Generalist Research Group, Division of General Pediatrics, and cDivision of Biostatistics, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania; dCenter for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; eLeonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. No consensus exists among general pediatricians or pediatric rheumatologists regarding whether corticosteroid therapy ameliorates the acute manifestations of Henoch-Schönlein purpura or mitigates renal injury. Therefore, we sought to synthesize the reported experimental and observational data regarding corticosteroid use.

METHODS. We performed a meta-analysis based on a comprehensive review of the literature in the Medline database (1956 to January 2007) and the Cochrane Controlled Trials Register. On the basis of reported outcomes among patients with Henoch-Schönlein purpura who were treated at diagnosis with corticosteroids compared with patients treated with supportive care only, we calculated odds ratios for the resolution of abdominal pain, the need for surgical intervention secondary to severe pain or intussusception, the likelihood of Henoch-Schönlein purpura recurrence, and the development of transient or persistent renal disease.

RESULTS. Of 201 articles retrieved from the initial literature search, 15 were eligible for inclusion. Corticosteroid treatment did not reduce the median time to resolution of abdominal pain but did significantly reduce the mean resolution time and increased the odds of resolution within 24 hours. Early corticosteroid treatment significantly reduced the odds of developing persistent renal disease. In addition, although the results were not statistically significant, the prospective data suggest reduced odds of both surgical intervention and recurrence.

CONCLUSIONS. Corticosteroids, given early in the course of illness, seem to produce consistent benefits for several major clinically relevant Henoch-Schönlein purpura outcomes.

www.pediatrics.org/cgi/doi/10.1542/peds.2007-0667
doi:10.1542/peds.2007-0667

Key Words
Henoch-Schönlein purpura, corticosteroids, children, meta-analytic methods, systematic reviews

Abbreviations
HSP—Henoch-Schönlein purpura
ESRD—end-stage renal disease
CCTR—Cochrane Controlled Trials Register
OR—odds ratio
CI—confidence interval

Accepted for publication May 21, 2007
Address correspondence to Pamela F. Weiss, MD, Division of Rheumatology, Children’s Hospital of Philadelphia, Children’s Seashore House Room 236, 34th and Civic Center Boulevard, Philadelphia, PA 19104-4399. E-mail: weisspa@email.chop.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
HENoch-Schönlein purpura (HSP) is the most common vasculitis of childhood, affecting between 8 and 20 per 100,000 children annually and accounting for 49% of all childhood vasculitides in the United States. Although typically self-limited, HSP can cause gastrointestinal hemorrhage, intussusception, and end-stage renal disease (ESRD). Renal involvement, manifested by hematuria, proteinuria, nephrotic syndrome, or renal insufficiency, may occur in 60% of children. In 1 recent study, 54% of patients developed renal manifestations within 3 months of diagnosis, 11.6% had persistent abnormalities after 7 years, and none developed ESRD. However, previous studies have reported that as much as 21% of children with HSP-associated nephritis develop rapidly progressive glomerulonephritis; 15% of these children with HSP-associated rapidly progressive glomerulonephritis, which is to say 2% of all patients with HSP, may progress to have uremia or ESRD. Furthermore, even seemingly benign HSP may have long-lasting effects: 2 studies have reported that 40% and 70% of term pregnancies of women with a history of childhood-onset HSP were complicated by hypertension, proteinuria, or preeclampsia.

The goals of treating HSP are typically to (1) ameliorate acute symptoms, (2) mitigate short-term morbidity (such as abdominal complications that require surgery), and (3) prevent chronic renal insufficiency. Because HSP is characterized by leukocyte infiltration of the blood vessel walls along with immunoglobulin A deposition (with resulting vascular injury and necrosis), and because corticosteroids inhibit inflammatory processes, early treatment with corticosteroids has been postulated to be effective for all 3 therapeutic goals, but much controversy remains. Although the literature is replete with retrospective studies that evaluated corticosteroid use for HSP, currently there are only 3 published prospective, placebo-controlled studies on the subject, and they vary in their conclusions regarding the utility of early corticosteroid administration.

Through a systematic review and meta-analysis, we sought to compare and contrast the experimental and observational data regarding corticosteroid use and, where appropriate, synthesize the data for 5 main clinical questions: (1) Do corticosteroids shorten the duration of abdominal symptoms in HSP? (2) Do corticosteroids decrease the odds of surgical intervention for HSP? (3) Do corticosteroids decrease the odds of disease recurrence? (4) Do corticosteroids decrease the odds of renal disease (transient plus persistent) in HSP? and (5) Do corticosteroids decrease the odds of developing persistent renal disease in HSP?

**METHODS**

**Data Sources**

The published medical literature was searched by using the Cochrane Controlled Trials Register (CCTR) and Medline database. Articles written in all languages were included in the search, and translation was obtained when needed.

CCTR and Medline were searched by using the Medical Subject Headings (MeSH) terms “steroids,” “methylprednisolone,” and “dexamethasone” and the key words “Henoch” and “corticosteroids.” “Henoch” and “purpura, Schönlein-Henoch” were grouped together and joined by “or.” “Steroids,” “methylprednisolone,” “dexamethasone,” and “corticosteroids” were grouped together and joined by “or.” Both groups of terms were joined together by “and.” The only limit applied to the search was “all children: 0–18 years.”

Given the large number of studies obtained from the literature search (N = 201), we cite in this report only the studies that were chosen for inclusion. A full list of publications is available at www.pediatric-generalists.org/weiss.htm.

**Study Eligibility**

Eligible studies were limited to those that examined the use of corticosteroids for the treatment of HSP; observational and randomized, controlled trials were included. An article was excluded if it (1) was a review, (2) examined therapy with a drug other than corticosteroids, (3) was a case report with fewer than 5 subjects with HSP, (4) focused on individuals older than 18 years, (5) did not discuss therapy with corticosteroids, (6) included only patients with nephritis at study onset, (7) did not assess definite outcomes, or (8) did not discuss HSP.

**Study Selection**

The initial literature search of the CCTR and Medline databases yielded 201 articles in 14 languages. Titles were reviewed to screen for eligibility. If the title yielded insufficient data to determine if a study was eligible, the abstract was obtained and reviewed. Articles without sufficient information in the abstract or those without an accessible abstract were examined in full text. Interpreters who were familiar with medical language and study designs evaluated all articles written in languages other than English.

Two of the authors (Drs Weiss and Feinstein) independently screened each of the potential titles, abstracts, and articles to determine inclusion. Disagreements were resolved by discussion and consensus mediated by a third author (Dr Feudtner). Reasons for article exclusion are presented in Table 1. After all the exclusions were applied, 15 articles remained for further analysis (Table 2).

In an attempt to find all relevant articles, the reference lists of included articles were searched and yielded 2 additional articles for potential inclusion. However, neither article was included because the outcomes could not be abstracted. In addition, the authors of the in-
included articles were contacted to learn of any unpublished trials or studies.

Data Extraction
Two authors (Drs Weiss and Feinstein) independently abstracted data from the remaining articles. Disagreements were resolved by discussion and consensus mediated by a third author (Dr Feudtner). Abstracted data from the remaining articles included information regarding the study population, existence of a control group, study limitations, and information relating to the resolution of abdominal pain, surgical intervention, recurrence, and the incidence of renal sequelae.

Statistical Analyses
Pooled odds ratios (ORs) were obtained by using the “metan” command for Stata 9.2 (Stata Corp, College Station, TX) for each of 5 clinical outcomes: (1) resolution of abdominal pain; (2) surgical intervention for severe abdominal pain or intussusception; (3) recurrence; (4) cumulative renal abnormalities; and (5) persistent renal abnormalities. The odds of having each outcome in patients who were treated with corticosteroids was compared with patients treated with routine supportive care for both prospective and retrospective studies. Results from fixed-effects models are reported. Tests for heterogeneity were performed for each analysis as a way to evaluate to what extent the results were consistent and suitable for fixed-effect modeling. If the test for heterogeneity was significant, we did not calculate a pooled OR. The Egger test for bias was performed and funnel plots were performed for each of the clinical outcomes to evaluate for publication bias. Each of the funnel plots was symmetrical, which suggests the absence of publication bias. The Egger test for bias could only be applied to analyses with 2 or more studies.

RESULTS

Do Corticosteroids Shorten the Duration of Abdominal Symptoms in HSP?
Although there is no consensus regarding the indication for corticosteroid use in HSP, anecdotally, abdominal pain is the most common reason that corticosteroids are prescribed. Huber et al used 2 mg/kg per day of corticosteroids for 1 week, with weaning over week 2, whereas Ronkainen et al used 1 mg/kg per day of corticosteroids for 2 weeks, with weaning over weeks 3 and 4 (Table 2). Using median values of total abdominal pain duration, Huber et al did not find any difference between the treatment and control groups ($P = .8$) (Fig 1A). In contrast, Ronkainen et al found a mean reduction of 1.2 days of pain in patients who were treated with corticosteroids ($P = .03$) (Fig 1A). Three retrospective studies reported on such effects within the first 24 hours after corticosteroid administration. However, after analysis, there was evidence of heterogeneity among the included studies ($P = .015$). A likely source of heterogeneity was the timing of corticosteroid administration in relation to disease onset. In the study by Reinehr et al, 31 of 57 patients received corticosteroids >21 days after the onset of HSP, whereas in the other 2 studies the treatment group received corticosteroids within several days of diagnosis. Analysis excluding the study by Reinehr et al showed a statistically significant positive corticosteroid effect, and there was no significant heterogeneity (OR: 5.42; 95% confidence interval [CI]: 1.60–18.29; $P = .476$) (Fig 1B).

A study by Chao et al looked at a unique group of patients with HSP who had hepatobiliary involvement. Involvement was defined by elevated liver transaminase levels and abnormal abdominal ultrasound results. Of those treated with corticosteroids, 90% with hepatic involvement, 100% with cholecystitis, and 100% with gallbladder sludge recovered within 7 days. In contrast, in the untreated group, 25% with hepatic involvement, none of those with cholecystitis, and only 50% with gallbladder sludge had resolution within 7 days. All patients in both groups, however, had resolution of symptoms by 14 days.

Do Corticosteroids Decrease the Incidence of Surgical Intervention for HSP?
Only 3 studies (1 randomized, controlled trial and 2 observational trials) reported intussusception, a rare and potentially life-threatening abdominal complication during the acute phase of HSP. Huber et al provided the only prospective study that reported incidence of intussusception; the risk of intussusception was reduced, but not to a significant degree, in the group that received corticosteroids (OR: 0.16; 95% CI: 0.01–3.62). The 2 retrospective studies together suggest a protective effect of corticosteroid exposure (OR: 0.75; 95% CI: 0.13–4.46) (Fig 2). There was no evidence of heterogeneity between the studies ($P = .39$).
Does Early Treatment With Corticosteroids Decrease the Odds of HSP Recurrence?
Recurrences affect up to one third of children with HSP.13,21 Many of these children require additional hospital admissions and pharmacotherapy. The 2 prospective studies with recurrence data suggest a protective effect of corticosteroids (OR: 0.32; 95% CI: 0.07–1.49), and there was no significant heterogeneity (P = .70) (Fig 3).11,14 The 5 retrospective observational studies that examined HSP recurrence exhibited heterogeneity (P < .01), so a pooled OR is not reported.11,13,14,20–23 The ORs and 95% CIs are reported for each study (Fig 3). No evidence of publication bias was found (Egger test: P = .55).

Do Corticosteroids Decrease the Likelihood of Developing Cumulative Renal Abnormalities With HSP?
Eight studies have reported data on cumulative (transient or persistent) renal abnormalities during the year...
after diagnosis\textsuperscript{9,14,17,24}. Heterogeneity was present among both the prospective and retrospective studies ($P = .03$ and $P < .01$, respectively), so pooled ORs are not reported. Instead, the individual ORs and 95% CIs of each study, prospective and retrospective, are shown (Fig 4). No evidence of publication bias was found (Egger test: $P = .99$).

**Do Corticosteroids Decrease the Likelihood of Developing Persistent Renal Abnormalities?**

In the 3 prospective studies, early corticosteroid treatment significantly reduced the odds of developing persistent renal disease (OR: 0.43; 95% CI: 0.19–0.96) (Fig 5A).\textsuperscript{8,10,13} There was no evidence of marked heterogeneity among the studies ($P = .341$). In addition, no evidence of publication bias was found (Egger test: $P = .99$). The retrospective study by Saulsbury\textsuperscript{12} did not show a statistically significant difference in renal outcome between the exposed and unexposed patients (OR: 1.25; 95% CI: 0.29–5.37) (Fig 5B).

### Sensitivity Analyses

**Does the Dose of Corticosteroid Matter?**

We investigated if the corticosteroid dose affects the likelihood of developing persistent renal disease by using unrestricted and restricted regression models and the
likelihood-ratio test. The doses and duration of treatment for each study are listed in Table 2. Two studies were excluded from this analysis because corticosteroid dose was not reported.20,23 The result of the likelihood-ratio test was statistically insignificant (P = .07), which suggests but fails to identify a significant dose-response effect.

What Future Study Would Reverse the Findings of This Meta-analysis?

Sensitivity analyses were performed to determine what sample size and magnitude of effect of a future hypothetical study would be required to reverse the findings, from the prospective studies, regarding the benefit of corticosteroid to decrease persistent renal disease presented in this meta-analysis. We conducted sensitivity analyses with hypothetical sample sizes of 200 and 400 patients and assuming an incidence of 5% and 20% persistent renal involvement in the control group (Fig 6). If the baseline risk of renal involvement in controls is 5%, a new study with 200 patients would need an effect size greater than an OR of 3.0 for the pooled OR to reach
1.0. A study with 400 patients would need an effect size greater than an OR of 2.25 for the new pooled OR to reach 1.0; for the new pooled OR to be statistically significant, the effect size would need to be larger than 3 (Fig 6A). If the baseline risk of renal involvement in controls is 20%, the OR of a new study with 200 patients would need to be >1.75 to raise the pooled OR to 1.0 and >3.0 to achieve significance. A new study with 400 patients would need an OR of >1.4 to raise the pooled OR to 1.0 and >2.25 to achieve statistical significance (Fig 6B).

**DISCUSSION**

This systematic review and data summary of 15 eligible articles suggests that early treatment of corticosteroids for children with HSP is associated with statistically significant increased odds of abdominal pain resolution within 24 hours and reduced odds of persistent renal disease. In addition, although the analyses lacked sufficient statistical precision, the likelihood of surgical intervention and of HSP recurrence may also be reduced. Overall, across all analyses, the pattern of effect is in the direction favoring the use of corticosteroids; none of the analyses indicated harm.

This review emphasizes the necessity for a more complete understanding of the ways in which corticosteroids impact the course of HSP, in both the acute and chronic settings. Corticosteroids were first postulated to benefit children with HSP in the 1950s and are effective in the treatment of other vasculitides in children and adults. Corticosteroids are the cornerstone of treatment for the majority of juvenile vasculitides including systemic lupus erythematosus, Wegener granulomatosis, polyarteritis nodosa, and Takayasu arteritis. By contrast, corticosteroid treatment is controversial for use for Kawasaki disease, another acute childhood vasculitis with potential morbidity and mortality; 1 recent report found no effect of corticosteroid, whereas others have heralded the beneficial effects of corticosteroids. Given this information should corticosteroid be given to children who present to the hospital with new-onset HSP? A rigorous answer to this question would have substantial clinical implications. Despite the widespread use of corticosteroids for vasculitides, there is no consensus yet among physicians as to whether corticosteroids should be given for HSP and, if so, for what indications.

There are several important potential limitations of the evidence we present. First, the definition of renal involvement differed among the studies, which means that different dimensions or degrees of HSP severity were potentially captured. For example, among the 3 prospective studies, the definition of proteinuria ranged from 200 to 400 mg/L, and the definition of hematuria ranged from 5 to 10 red blood cells per high-powered field. Second, the dosing regimens and mode of delivery for corticosteroids were not the same in each study (Table 2). When corticosteroid therapy was stratified by high- versus low-dose steroids for each of the clinical
outcomes, there were no statistical differences between the 2 groups (data not shown). Although various doses of corticosteroids or the method of delivery (oral versus intravenous) may affect the size of the risk reduction, these changes in dose or route are unlikely to reverse the direction of effects. Third, the studies did not uniformly classify HSP cases as incident or recurrent. Therefore, the potential exists to inappropriately group severe recurrence with milder recurrence such as rash alone. Likewise, given uncertainty about loss to follow-up, particularly in the corticosteroid-treated group, the protective effect of corticosteroids is likely to be an underestimate.

Two other limitations should be kept in mind when interpreting these data. First, our ability to draw robust conclusions was hampered by the limited number of prospective studies and small numbers of patients within the studies, which often resulted in imprecise measures of treatment effects. Although the effect of corticosteroid on several of the outcomes did achieve statistically significant evidence of benefit, the failure to demonstrate benefit regarding other outcomes and the failure to show a significant dose-response effect may reflect either the still relatively underpowered nature of the combined existing studies or a true absence of impact. Second, the effects noted in the retrospective observational studies probably include some degree of confounding by indication, whereby patients with greater disease severity were more likely to receive corticosteroids. Given that most of the retrospective studies nevertheless continued to show a benefit, the results across all studies are more reassuring.

These caveats notwithstanding, our findings suggest that the potential benefit of corticosteroid administration early in the course of HSP may be more prominent than previously suggested for both acute (pain, surgical intervention) and chronic (recurrence, renal disease) complications of disease. If corticosteroid therapy is truly as effective as this meta-analysis suggests, then broader use of corticosteroids is likely to be an underestimate.

Two other limitations should be kept in mind when interpreting these data. First, our ability to draw robust conclusions was hampered by the limited number of prospective studies and small numbers of patients within the studies, which often resulted in imprecise measures of treatment effects. Although the effect of corticosteroid on several of the outcomes did achieve statistically significant evidence of benefit, the failure to demonstrate benefit regarding other outcomes and the failure to show a significant dose-response effect may reflect either the still relatively underpowered nature of the combined existing studies or a true absence of impact. Second, the effects noted in the retrospective observational studies probably include some degree of confounding by indication, whereby patients with greater disease severity were more likely to receive corticosteroids. Given that most of the retrospective studies nevertheless continued to show a benefit, the results across all studies are more reassuring.

These caveats notwithstanding, our findings suggest that the potential benefit of corticosteroid administration early in the course of HSP may be more prominent than previously suggested for both acute (pain, surgical intervention) and chronic (recurrence, renal disease) complications of disease. If corticosteroid therapy is truly as effective as this meta-analysis suggests, then broader use of corticosteroids is likely to be an underestimate.

Two other limitations should be kept in mind when interpreting these data. First, our ability to draw robust conclusions was hampered by the limited number of prospective studies and small numbers of patients within the studies, which often resulted in imprecise measures of treatment effects. Although the effect of corticosteroid on several of the outcomes did achieve statistically significant evidence of benefit, the failure to demonstrate benefit regarding other outcomes and the failure to show a significant dose-response effect may reflect either the still relatively underpowered nature of the combined existing studies or a true absence of impact. Second, the effects noted in the retrospective observational studies probably include some degree of confounding by indication, whereby patients with greater disease severity were more likely to receive corticosteroids. Given that most of the retrospective studies nevertheless continued to show a benefit, the results across all studies are more reassuring.

These caveats notwithstanding, our findings suggest that the potential benefit of corticosteroid administration early in the course of HSP may be more prominent than previously suggested for both acute (pain, surgical intervention) and chronic (recurrence, renal disease) complications of disease. If corticosteroid therapy is truly as effective as this meta-analysis suggests, then broader use of corticosteroids is likely to be an underestimate.

DEBATE ON ENDING SAT GAINS GROUND

“The social scientist Charles Murray has a knack for noisily tapping into cultural preoccupations. In his 1984 book, ‘Losing Ground,’ he argued that welfare perpetuated dependency and should be eliminated. In *The Bell Curve* (1994), which he wrote with Richard J. Herrnstein, he argued that those who get ahead in America (mostly whites) are genetically endowed with more intelligence than those who do not (disproportionately African-Americans). Now Mr. Murray is at it again, proposing in a recent article to abolish the SAT. This position cannot help but provoke a double-take. After all, while making his arguments about genes, race and intelligence, Mr. Murray promoted the IQ test as a reliable measure of aptitude. Yet he is suggesting that one of the most widely used assessment tests be eliminated. With so many college officials and parents dissatisfied with the SAT, even those who think Mr. Murray’s other theories are misguided or offensive could find themselves agreeing with him on this issue. Unlike other critics of the SAT, Mr. Murray does not see the test as flawed, nor does he think that the wealthy have an unfair advantage because they can buy expensive coaching. But he recognizes that most people do not agree with him and believe the test is rigged to favor the rich. ‘It is a corrosive symbol of privilege,’ he said. And so, he concludes that college admissions offices should reject the SAT and substitute other standardized tests: subject or so-called achievement tests that gauge knowledge in specific disciplines like history or chemistry. ‘This is really a hot topic,’ said William R. Fitzsimmons, the dean of admissions and financial aid at Harvard University.”


Noted by JFL
Inhaled Nitric Oxide for Preterm Infants: A Systematic Review

Keith J. Barrington, MB, ChB, Neil N. Finer, MD

Department of Pediatrics, McGill University, Montreal, Quebec, Canada; Department of Pediatrics, University of California, San Diego, California

Financial Disclosure: Dr Barrington was chairperson of the Canadian Medical Advisory Committee for iNO Therapeutics for a meeting in 2004; Dr Finer has indicated he has no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. Our goal was to determine whether, for preterm newborn infants with respiratory disease, inhaled nitric oxide reduced the rates of death, bronchopulmonary dysplasia, intracranial hemorrhage, or neurodevelopmental disability.

METHODS. We searched Medline, Embase, Healthstar, and the Cochrane Central Register of Controlled Trials using the search terms “nitric oxide,” “clinical trial,” and “newborn” and covering 1985–2006. We also searched abstracts of the Pediatric Academic Societies.

RESULTS. Eleven randomized, controlled trials of inhaled nitric oxide therapy for preterm infants were found. The trials were grouped into 3 categories according to the entry criteria, that is, entry in the first 3 days of life on the basis of oxygenation criteria (early rescue), enrollment after 3 days on the basis of elevated risk of bronchopulmonary dysplasia, and routine use for intubated preterm infants. Early rescue treatment based on oxygenation criteria did not seem to affect mortality or bronchopulmonary dysplasia rates. Routine use for intubated preterm infants showed a barely significant reduction in the incidence of the combined outcome of death or bronchopulmonary dysplasia (relative risk [RR]: 0.91 [95% confidence limits (CLs): 0.84, 0.99]). Later treatment based on the risk of bronchopulmonary dysplasia showed no significant effect on this outcome. Early rescue treatment showed a trend toward increased incidence of severe intracranial hemorrhage, whereas routine use for intubated preterm infants seemed to show a reduction in the incidence of either severe intracranial hemorrhage or periventricular leukomalacia (RR: 0.70 [95% CLs: 0.53, 0.91]).

CONCLUSIONS. Inhaled nitric oxide as rescue therapy for very ill preterm infants undergoing ventilation does not seem to be effective and may increase severe intracranial hemorrhage. Later use of inhaled nitric oxide to prevent bronchopulmonary dysplasia does not seem to be effective. Early routine use of inhaled nitric oxide for mildly sick, preterm infants seems to decrease the risk of serious brain injury and may improve rates of survival without bronchopulmonary dysplasia.
SINCE THE INTRODUCTION of surfactant therapy, mortality rates for preterm infants have decreased significantly.\textsuperscript{1} Occasionally infants do not experience adequate improvement in oxygenation after surfactant treatment, however, and other complications of prematurity continue to cause substantial long-term disability.\textsuperscript{2}

Premature animals with models of hyaline membrane disease have elevated pulmonary vascular resistance, and both pulmonary artery pressure and oxygenation may be improved with inhaled nitric oxide (iNO).\textsuperscript{3,4} Preterm infants with respiratory failure also have increased pulmonary artery pressure.\textsuperscript{5} This is rarely sufficient to cause reversal of ductal shunts, however, and the hemodynamic profile thus differs from that of term neonates with severe pulmonary hypertension. In term neonates with hypoxic respiratory failure, iNO decreases the requirement for extracorporeal membrane oxygenation without decreasing mortality rates.\textsuperscript{6} However, because of different pathophysiologic features, different entry criteria, and different outcomes, the results for term infants cannot be extrapolated to preterm infants.

Preterm infants are at risk of long-term pulmonary disability attributable to bronchopulmonary dysplasia (BPD). If iNO therapy leads to a decrease in required ventilatory support, then reductions in lung injury and the frequency of BPD may follow. BPD is important because it leads to chronic medical illness and rehospitalization and is associated with neurodevelopmental impairment.\textsuperscript{7,8} However, nitric oxide has both prooxidant and antioxidant activities and can potentially worsen lung injury.\textsuperscript{9} Therefore, the effects of iNO therapy on developing lungs must be evaluated carefully before the introduction of such therapy into clinical practice.

Of particular concern for preterm infants is the fact that iNO affects coagulation.\textsuperscript{10,11} Preterm infants are at high risk of developing intracranial hemorrhage, which has substantial effects on long-term developmental outcomes. Therefore, it is important that iNO be evaluated for its effect on intracranial hemorrhage in preterm infants.

The few case reports and case series published before randomized, controlled trials were conducted demonstrated that premature infants with severe respiratory failure that did not respond to full conventional management, including surfactant therapy and high-frequency ventilation, might experience improved oxygenation with iNO.\textsuperscript{12,13} In those reports, death and intracranial hemorrhage were frequent.

The objective of this study was to determine whether, in preterm newborn infants with respiratory disease, treatment with iNO improves oxygenation and reduces the rates of death, BPD, intracranial hemorrhage or other serious brain injury, and adverse, long-term, neurodevelopmental outcomes. This is an edited version of an updated systematic review to be published in the Cochrane Database of Systematic Reviews.

**METHODS**

The protocol for the systematic review was approved by the Cochrane Neonatal Review Group. The criteria for considering studies for this review were as follows: the studies were randomized or quasi-randomized clinical trials, and the participants were preterm infants (<35 weeks’ gestation) with respiratory failure after adequate treatment with surfactant. The intervention was iNO therapy, compared with placebo or control treatment, in addition to conventional treatment, for respiratory failure. The outcome measures in which we were interested were death before hospital discharge, BPD (oxygen dependence at corrected age of 36 weeks), death or BPD (at corrected age of 36 weeks), intracranial hemorrhage or severe intracranial hemorrhage, periventricular leukomalacia, neurodevelopmental disability (proportion of survivors with neurologic abnormalities sufficient to affect quality of life, developmental index $>2$ SDs below the mean, using a validated scale at $>12$ months of age, or mean developmental index among survivors at $>12$ months of age), severe retinopathy of prematurity, and oxygenation within 2 hours after therapy.

The following search strategy was used for identification of studies. A Medline search was performed by using PubMed and the search terms “nitric oxide” and “newborn.” This search was limited to clinical trials and was updated most recently in November 2006. The years searched were 1985 to the present. The abstracts of the annual Pediatric Academic Societies meetings were also searched from 1995 to 2006. In addition, the standard methods of the Cochrane Neonatal Review Group were used to find any additional references; this involved searches of Embase, Healthstar, and the Cochrane Central Register of Controlled Trials, all last updated in November 2006. Each identified trial was assessed for methodologic quality with respect to (1) masking of allocation, (2) masking of intervention, (3) completeness of follow-up monitoring, and (4) masking of outcome assessment.

For categorical outcomes, typical estimates for relative risk (RR) and risk difference (RD) were calculated by using RevMan software (Nordic Cochrane Centre, Cochrane Collaboration, 2003, Copenhagen, Denmark), and 95% confidence limits (CLs) were used. A fixed-effect model was assumed. Continuous outcomes were analyzed by using weighted mean differences, also assuming a fixed-effect model. Heterogeneity was evaluated by using the $\chi^2$ statistic, which is reported whenever the result was $>50\%$; a $\chi^2$ test for heterogeneity was also performed, and results are reported for $P < .05$. Sensitivity analyses were performed within groups of trials when the trials were of very different design or quality.
Results both with and without inclusion of particular trials are presented.

RESULTS

Description of Studies

We found 11 published, randomized, controlled trials of iNO, compared with control treatment, for preterm infants. Abbreviated details are found in Table 1. All studies used an intent-to-treat approach to analysis.

The entry criteria for the 11 studies were quite dissimilar, with enrollment in the first 2 days of life for most of the infants in 7 of the trials, all of which required significantly impaired oxygenation for trial enrollment. Two of the remaining 4 trials had enrollment after the first 3 days of life for infants with elevated risks of developing BPD. The remaining trials, by Schreiber et al and Kinsella et al, enrolled infants in the first 2 days of life, and any preterm infant undergoing ventilation was considered eligible. The implications for clinical practice are clearly quite different for these 3 groups of studies, which were therefore analyzed separately.

Studies were eligible to be considered in the first group if they included acutely ill, preterm infants undergoing ventilation, most of whom were enrolled within the first 3 days of life, and the infants needed to satisfy a criterion regarding severity of illness; this group is termed “early rescue treatment.” The average oxygenation indices (OIs), when available, were as follows: median of 32 for the Neonatal Ventilation With Inhaled Nitric Oxide Versus Ventilatory Support Without Inhaled Nitric Oxide for Severe Respiratory Failure (INNOVO) trial, 10.8 to 11.9 for the trial by Srisuparp et al, 15 and only mild disease and required increasing OIs for enrollment of infants with increasing birth weight. Allocation was through a “card-picking scheme”; therefore, masking was uncertain. This was a pilot study for the study by Schreiber et al and therefore was underpowered. Baseline characteristics were similar between the groups.

Studies were included in the second group if they enrolled preterm infants who were >3 days of age and qualified because of an elevated risk for BPD. There were 2 such trials. Ballard et al reported a median respiratory severity score (cm H2O; which is not directly convertible to the OI because it does not take into account the PaO2; few eligible infants would have had arterial lines), calculated as the fraction of inspired oxygen (FIO2) multiplied by the mean airway pressure (in centimeters of water), of 3.5 in both the treated and control infants, which suggests mild illness.

Studies were eligible for the third group if they enrolled infants early in life (<3 days of age) and had no criteria regarding severity of illness, other than being intubated. In the study by Kinsella et al, the mean OIs were 5.4 (iNO group) vs 5.8 (control group); in the study by Schreiber et al, the median OI was 6.94. This group is termed “early routine use.”

Trials of Early Rescue Treatment With iNO for Preterm Infants

Kinsella et al randomly assigned preterm infants if the ratio between PaO2 and alveolar oxygen pressure (calculated as [FIO2 × (atmospheric pressure – 47)] – PaCO2) was <0.1 in 2 successive blood gas analyses in the first 7 days of life (expected mortality rate: 50%). The planned sample size was 210 infants. A planned interim analysis, performed after 2.5 years, found no detectable difference in the main outcome (survival to discharge). Forty-eight infants had received iNO, and 32 were control subjects. The study was terminated at that time because the analysis suggested that significant benefit was unlikely to be detected. Baseline characteristics of the groups were similar, apart from a greater number of iNO-treated infants having no intracranial hemorrhage at the start of the study (73% vs 59%). The PaO2/FIO2 values were 42 (SD: 18) for the iNO group and 42 (SD: 16) for the control group.

The Franco-Belgium collaborative NO trial group performed a multicenter international trial. All except 1 of the 85 preterm infants (<33 weeks’ gestation) received surfactant, and the majority (75%) were receiving high-frequency ventilation. The majority of the infants were enrolled on the first or second day of life. The study was terminated early because of slowing enrollment after 27 months. After assessment of the primary outcome at 2 hours, treatment with open-label iNO was allowed if the infant’s OI exceeded 30. Five of the control infants eventually received iNO. The availability of backup treatment with iNO for control infants limited the ability of the study to address long-term outcomes. All of the baseline characteristics were similar between the groups.

Srisuparp et al performed a single-institution study that randomly assigned intubated preterm infants of <2000 g who had received surfactant and had clinical respiratory distress syndrome. OI thresholds, noted in Table 1, allowed random assignment of the smallest infants (<1000 g) with only mild disease and required increasing OIs for enrollment of infants with increasing birth weight. Allocation was through a “card-picking scheme”; therefore, masking was uncertain. This was a pilot study for the study by Schreiber et al and therefore was underpowered. Baseline characteristics were similar between the groups.

Hascoet et al conducted a study in 10 European centers. Any intubated preterm infant could be enrolled and assigned randomly, but the randomization was revealed only if the infants developed hypoxic respiratory failure, which led to 61 iNO-treated infants and 84 control infants actually being exposed to the study intervention. If an infant developed refractory hypoxemia at any time, then the case was defined as a failure and iNO was administered according to French Drug Agency recommendations. Refractory hypoxemia before 6 hours of age occurred in 20 infants, who therefore were not entered...
## TABLE 1  
**Studies Analyzed**

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Study Characteristics and Quality Indicators</th>
<th>Participants and Entry Criteria</th>
<th>Interventions, Dose, and Duration</th>
<th>Primary Outcomes</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early rescue treatment studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kinsella et al(^a) (1999)</td>
<td>Multicenter trial; masking of allocation; masking of intervention; complete follow-up data; masking of outcome assessment</td>
<td>80 preterm infants (≤34 wk gestation) (&lt;7 d of age, with PaO(_2)/alveolar oxygen pressure ratio of &lt;0.1 in 2 blood gas analyses after surfactant treatment)</td>
<td>iNO at 5 ppm (n = 48) or placebo (n = 52) for 7 d, after which “trials off” were allowed; maximal treatment duration: 14 d</td>
<td>Survival</td>
<td>Initially planned for 210 subjects</td>
</tr>
<tr>
<td>INNOVO trial(^b) (2005)</td>
<td>Multicenter trial; masking of allocation; no masking of intervention; complete follow-up data; no masking of outcome assessment</td>
<td>420 preterm infants (&lt;34 wk gestation) with OI of ≥10 in 2 blood gas analyses 30 min to 12 h apart, ≥4 h after surfactant treatment, 4–120 h after birth</td>
<td>iNO, initially at 5 ppm to 10 ppm if no response; treatment started after 2 h; maximal duration: 336 h</td>
<td>Death or BPD at PMA of 36 wk</td>
<td>Initially planned for 220 infants per arm; stopped by DSMC because of apparent increase in severe IVH with no evidence of benefit</td>
</tr>
<tr>
<td>Van Meurs et al(^c) (2005)</td>
<td>Multicenter trial; masking of allocation; no masking of intervention; complete follow-up data; no masking of outcome assessment</td>
<td>108 preterm infants (&lt;34 wk gestation) &lt;28 d of age, with “severe respiratory failure”</td>
<td>iNO usually at 5 ppm, could be increased up to 40 ppm (n = 55) or no supplemental gas (n = 53); maximal duration unclear; one half received iNO for &lt;3 d</td>
<td>Death or severe disability at corrected postnatal age of 1 y, or death or continued oxygen need at expected date of birth</td>
<td>Initially planned for 200 subjects</td>
</tr>
<tr>
<td>Hascoet et al(^d) (2004)</td>
<td>Multicenter trial; masking of allocation; no masking of intervention; complete follow-up data; no masking of outcome assessment</td>
<td>860 infants (&lt;32 wk gestation) enrolled, eligible for study gas if hypoxic respiratory failure (ie, need for mechanical ventilation, FiO(_2) of &gt;40%, and PaO(_2)/alveolar oxygen pressure ratio of &lt;0.22 at 6–48 h) occurred; number who met this criterion: 145 (iNO; 61; control, 84)</td>
<td>iNO at 5 ppm or standard care; increase to 10 ppm allowed if no response; weaning started after 2 h; maximal duration: 7 d</td>
<td>Severe IVH (grade 3 or 4)</td>
<td>Performed as preliminary pilot study, before study by Schreiber et al(^a)</td>
</tr>
<tr>
<td>Franco-Belgium trial group(^e) (1999)</td>
<td>Multicenter trial; masking of allocation; no masking of intervention; complete follow-up data; no masking of outcome assessment</td>
<td>85 preterm infants (&lt;33 wk gestation) with OI of 12.5–30 at &lt;7 d</td>
<td>iNO at 10 ppm (n = 40) or control (n = 45); open-label treatment with iNO allowed for control subjects with OI of &gt;30; maximal duration unclear</td>
<td>OI after 2 h of therapy</td>
<td>Initially planned for 360 infants across both gestational age strata</td>
</tr>
<tr>
<td>Srisuparp et al(^f) (2002)</td>
<td>Single-center trial; unclear masking of allocation; no masking of intervention; complete follow-up data; no masking of outcome assessment</td>
<td>34 infants of &lt;2000 g, undergoing ventilation after surfactant treatment, &lt;72 h of age, with OI of &gt;4 for birth weight of &lt;1000 g, &gt;6 for 1001–1250 g, &gt;8 for 1251–1500 g, &gt;10 for 1501–1750 g, and &gt;12 for 1751–2000 g</td>
<td>iNO at 20 ppm or standard trial; of weaning after 72 h; maximal duration: 7 d</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

\(^a\) Study of 145 premature infants treated with iNO and 84 controls. Some infants received 
different doses of iNO at different times. Therefore, we used two arms: iNO to all 
in the initial trial and iNO after the first trial. This trial was stopped by the DSMC 
because of an apparent increase in severe IVH. (1); open-label treatment with 
iNO allowed for control subjects with OI of >30; maximal duration unclear. 

\(^b\) Multicenter trial; masking of allocation; no masking of intervention; complete 
follow-up data; no masking of outcome assessment. 108 preterm infants (<34 wk 
gestation) <28 d of age, with “severe respiratory failure.”

\(^c\) Multicenter trial; masking of allocation; no masking of intervention; complete 
follow-up data; no masking of outcome assessment. 420 preterm infants (<34 wk 
gestation) with OI of ≥10 in 2 blood gas analyses 30 min to 12 h apart, ≥4 h 
after surfactant treatment, 4–120 h after birth.

\(^d\) Multicenter trial; masking of allocation; no masking of intervention; complete 
follow-up data; no masking of outcome assessment. 860 infants (<32 wk gestation) 
rolled, eligible for study gas if hypoxic respiratory failure (ie, need for mechanical 
ventilation, FiO\(_2\) of >40%, and PaO\(_2\)/alveolar oxygen pressure ratio of <0.22 at 6–48 h) 
occurred; number who met this criterion: 145 (iNO; 61; control, 84).

\(^e\) Multicenter trial; masking of allocation; no masking of intervention; complete 
follow-up data; no masking of outcome assessment. 85 preterm infants (<33 wk 
gestation) with OI of 12.5–30 at <7 d.

\(^f\) Single-center trial; unclear masking of allocation; no masking of intervention; 
complete follow-up data; no masking of outcome assessment. 34 infants of <2000 g, 
undergoing ventilation after surfactant treatment, <72 h of age, with OI of >4 for 
birth weight of <1000 g, >6 for 1001–1250 g, >8 for 1251–1500 g, >10 for 
1501–1750 g, and >12 for 1751–2000 g.
into the study. An additional 20 iNO-treated infants and 28 control infants received open-label iNO for treatment of refractory hypoxemia, which complicated the analyses of the results with significant contamination of the control group. The initially planned sample size was achieved.

Van Meurs et al\(^7\) performed a multicenter study in which initial entry criteria restricted enrollment to in-

### TABLE 1 (Continued)

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Study Characteristics and Quality Indicators</th>
<th>Participants and Entry Criteria</th>
<th>Interventions, Dose, and Duration</th>
<th>Primary Outcomes</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dani et al(^19) (2006)</td>
<td>Single-center trial; masking of allocation; no masking of intervention; complete follow-up data; no masking of outcome assessment</td>
<td>40 preterm infants (30 wk gestation) with severe RDS undergoing ventilation, with $F_{O_2}$ of &gt;0.5 and $P_{A_{O_2}}$/alveolar oxygen pressure ratio of &lt;0.15, despite surfactant treatment, at &lt;7 d of age</td>
<td>Either iNO at 10 ppm for 4 h and then 6 ppm or no treatment; weaning started at 72 h; iNO treatment stopped when infant was extubated or when $F_{O_2}$ was &lt;0.3 with mean airway pressure of &lt;8 cm H$_2$O</td>
<td>Death or BPD at PMA of 36 wk</td>
<td>Initially planned for 52 subjects; unplanned interim analysis performed because of impression that results were significant</td>
</tr>
<tr>
<td>Subhedar et al(^23) (1997)</td>
<td>Single-center trial; masking of allocation; no masking of intervention; complete follow-up data; no masking of outcome assessment</td>
<td>42 preterm infants (&lt;32 wk gestation) with “high risk” of developing BPD, at 96 h of age; comparing iNO, dexamethasone, combined therapy, and control treatment, using a 2 × 2 factorial design</td>
<td>iNO initially administered at 20 ppm and weaned if effective (n = 20) or control (n = 22); dexamethasone or no steroids (n = 21 per group); weaning or cessation of iNO therapy attempted at 72 h and every 24 h thereafter</td>
<td>Death or BPD at PMA of 36 wk</td>
<td>Initially planned for 88 subjects</td>
</tr>
<tr>
<td>Ballard et al(^24) (2006)</td>
<td>Multicenter trial; masking of allocation; masking of intervention; complete follow-up data; masking of outcome assessment</td>
<td>582 infants of &lt;1230 g receiving assisted ventilation at 7–21 d (or, if &lt;800 g, underlying CPAP)</td>
<td>iNO initially at 20 ppm for 48–96 h; doses subsequently decreased 10, 5, and 2 ppm at weekly intervals; minimal treatment duration: 24 d</td>
<td>Death or BPD at PMA of 36 wk</td>
<td></td>
</tr>
<tr>
<td>Schreiber et al(^21) (2003)</td>
<td>Single-center trial; masking of allocation; masking of treatment; complete follow-up data; masking of outcome assessment</td>
<td>207 infants (&lt;34 wk gestation), &lt;72 h of age, intubated and undergoing ventilation because of RDS, with birth weight of &lt;2000 g; 2 × 2 factorial design also investigating HFV versus conventional ventilation</td>
<td>iNO at 10 ppm for 12–24 h and then 5 ppm or placebo; total duration: 7 d</td>
<td>Death or BPD at PMA of 36 wk</td>
<td>Factorial design with HFV and iNO</td>
</tr>
<tr>
<td>Kinsella et al(^22) (2006)</td>
<td>Multicenter trial; masking of allocation; masking of intervention; complete follow-up data; masking of outcome assessment</td>
<td>793 preterm infants (&lt;34 wk gestation) with respiratory failure needing assisted ventilation in first 48 h and expected to need it for ≥48 h</td>
<td>iNO at 5 ppm (n = 398) or no iNO (n = 395) for 21 d or until extubation</td>
<td>Death or BPD</td>
<td>Baseline and follow-up cranial ultrasound scans were required</td>
</tr>
</tbody>
</table>

RDS indicates respiratory distress syndrome; HFOV, high-frequency ventilation; PMA, postmenstrual age; DSMC, data safety-monitoring committee; IVH, intraventricular hemorrhage; CPAP, continuous positive airway pressure.
fants with gestational ages of <34 weeks and birth weights of 401 to 1500 g who were undergoing assisted ventilation ≥4 hours after surfactant therapy and were considered at high risk because of OIs of ≥10 in 2 consecutive blood gas analyses. This criterion was revised after the initial interim analysis showed an unexpectedly high mortality rate with an OI of ≥5 followed by an OI of ≥7.5 after ≥30 minutes. Although it was initially planned to recruit 440 infants, the study was terminated after two thirds of the infants had been assessed for the primary outcome, because there seemed to be an increase in the severe intracranial hemorrhage rate but no benefit in the primary outcome. By the time the analysis was completed, 420 infants had been enrolled and the study was terminated. Baseline characteristics were similar between the groups.

The INNOVO trial16 was a European multicenter study with masked allocation using a telephone system; treatment assignment was through minimization25 “with a probabilistic element,” rather than strict randomization. Eligibility criteria were “severe respiratory failure requiring assisted ventilation if the responsible physician was uncertain about whether an infant might benefit from iNO.” iNO treatment was suggested to be initiated at 5 ppm and could be doubled up to a maximum of 40 ppm. There were 2 primary outcome criteria listed in the main publication, but the sample size was calculated on the basis of a reduction in the frequency of the combined outcome of death or severe disability at corrected postnatal age of 1 year. The study planned to enroll 200 infants at <34 weeks’ gestation and <28 days of age; 55 iNO-treated and 53 control infants were enrolled, and the study was stopped at the end of calendar year 2001, which was apparently preplanned. Follow-up data were complete for all except 1 infant and were not formally masked.

Dani et al19 performed a single-center study with infants who were undergoing ventilation and were experiencing severe respiratory distress despite surfactant treatment. The mean age of intervention initiation was 43 hours for the iNO group. This single-center study planned to enroll 52 infants but was terminated after 40 subjects had been enrolled, after a previously unplanned interim analysis that confirmed the investigators’ impression that BPD was less frequent in the iNO group. This early termination provided insufficient protection from type 1 errors.

Trials of Routine Use of iNO for Intubated Preterm Infants
Schreiber et al21 performed a single-center study that enrolled intubated preterm infants (gestational age of <34 weeks and birth weight of <2 kg) at <72 hours of age, without additional criteria regarding severity of illness. This was a 2 × 2 factorial design examining 7 days of iNO treatment, compared with oxygen placebo treatment; the second comparison was the use of high-frequency oscillatory ventilation versus conventional ventilation. The planned sample size was achieved, and follow-up data were complete. Assessment of the primary outcome was performed in a masked manner. It was not stated whether the long-term neurodevelopmental follow-up monitoring was also masked.

Kinsella et al22 performed the largest of the multicenter trials completed to date. Subjects were <34 weeks’ gestational age, were undergoing ventilation because of respiratory failure in the first 48 hours, and were expected to remain intubated for >48 hours. There were no additional requirements regarding severity of illness. The planned sample size of 792 infants was achieved. The study was overseen by a data safety-monitoring committee, with interim analyses according to unplanned rules. Groups were well balanced. The follow-up data were complete with respect to assessment of the primary outcome, which was assessed in a masked manner.
Analyses
The usefulness of overall analyses was considered to be limited, because of the differing entry criteria for the studies; the criteria regarding severity of illness and age at entry varied so greatly that pooling the results was not considered appropriate. Control group mortality rates also varied substantially (6%–64%), emphasizing the differences in the eligible patients. Therefore, we performed analyses only according to the aforementioned posthoc groupings.

Comparability of Studies Within Groups
Early Rescue Treatment Studies
The early rescue treatment studies all randomly assigned the infants to low-dose iNO or control treatment. The INNOVO study did not have a reproducible criterion for entry; despite this difference, the mortality and BPD rates were higher but not very dissimilar, compared with the other studies in the early rescue treatment group. The methods used for calculation of the oxygenation defect in the remaining studies were different, with some studies reporting OI and others reporting PaO2/FIO2, which cannot be compared directly. In the early rescue treatment studies, the majority of patients were enrolled before 3 days of age, although some studies allowed enrollment up to 7 days of age. Of note, both Hascoet et al and the Franco-Belgium group allowed backup treatment of control subjects with iNO if their condition worsened to a prespecified degree. This might lead to underestimates of both benefits and risks. For this reason, sensitivity analyses were performed by excluding those 2 studies.

The early rescue treatment studies had comparable mortality rates for the control groups, with (except for the smallest study) mortality rates of >30%. The highest control group mortality rate was in the INNOVO study (64%), with the rest being between 30% and 44%.

Studies With Later Entry Based on BPD Risk
Two studies evaluated infants >3 days of age on the basis of elevated risk of BPD; the studies were quite different from each other. Subhedar et al investigated both iNO and dexamethasone therapy, by using a factorial design, for infants with an extremely high risk of BPD (almost 100% among survivors). Ballard et al enrolled infants with increased BPD risk solely on the basis of the continued need for respiratory support. Because of these differences in trial design, we performed sensitivity analyses with and without the trial by Subhedar et al, which showed no major effect because that trial was very small. The mortality rate for the control group in the study by Ballard et al was only 6%, which reflects the older age at entry than in the early rescue and early routine treatment studies and the lesser severity of illness than in the early rescue treatment studies.

Early Routine Treatment Studies
Two studies enrolled infants without specific criteria for disease severity. Schreiber et al randomly assigned preterm infants who were ventilator-dependent after receiving surfactant, and Kinsella et al enrolled infants with gestational ages of <34 weeks who were expected to undergo ventilation for >48 hours. The infants in this group of studies had much lower OIs than did the infants in the early rescue treatment studies. The control group mortality rates were quite similar between these 2 studies (23% and 25%, respectively).

Outcomes

Death Before Hospital Discharge
All trials assessed survival to discharge, and none of the individual trials showed a significant effect. The early rescue treatment studies had a typical RR of 1.05 (95% CLs: 0.91, 1.22; RD: 0.02; 95% CLs: −0.04, 0.09). The studies with entry after 3 days on the basis of BPD risk had a RR of 1.06 (95% CLs: 0.64, 1.74; RD: 0.00; 95% CLs: −0.04, 0.05). For the studies of early routine treatment, the typical estimate of the RR for death before hospital discharge showed a significant decrease, with an upper 95% CL that approached 1 (typical RR: 0.77; 95% CLs: 0.60, 0.98; RD: −0.06; 95% CLs: −0.11, 0.01) (Fig 1).

Death Before Postmenstrual Age of 36 Weeks
Six studies reported this outcome, with early rescue treatment. There was no significant effect of iNO on this outcome (typical RR: 0.89; 95% CLs: 0.72, 1.11; RD: −0.05; 95% CLs: −0.13, 0.14). The study by Subhedar et al with entry after 3 days on the basis of BPD risk, also reported this result and did not show a significant effect.

BPD (Oxygen Dependence Among Survivors at Corrected Age of 36 Weeks)
All of the published studies except those by Hascoet et al and Srisuparp et al reported BPD rates at 36 weeks, and none of the individual trials found a significant effect. There was substantial heterogeneity for each group of studies, none of which showed a statistically significant effect.

For early rescue treatment studies, the typical RR was 0.89 (95% CLs: 0.76, 1.05; \(I^2 = 47.8\%\); RD: −0.05; 95% CLs: −0.12, 0.02). For studies with entry after 3 days on the basis of BPD risk, the typical RR was 0.89 (95% CLs: 0.78, 1.02; \(I^2 = 85.5\%\); RD: −0.07; 95% CLs: −0.15, 0.01). For studies of routine use for intubated preterm infants, the typical RR was 0.96 (95% CLs: 0.85, 1.08; \(I^2 = 64.4\%\); RD: −0.02; 95% CLs: −0.09, 0.04).

Death or BPD
Data on the combined outcome of death or BPD (or its converse, survival without BPD) were available for all of the studies (Fig 2). None of the individual early rescue
treatment trials found a significant effect, and this group of studies showed no effect (typical RR: 0.95; 95% CLs: 0.88, 1.02: RD: −0.04; 95% CLs: −0.09, 0.02). Similarly, the studies with entry after 3 days on the basis of BPD risk did not show a significant effect individually, and the group results were not significant (typical RR: 0.90; 95% CLs: 0.80, 1.02: RD: −0.06; 95% CLs: −0.14, 0.01).

The studies of routine use for intubated preterm neonates showed a significant reduction; of note, the upper 95% CL was 0.99 (RR: 0.91; 95% CLs: 0.84, 0.99). The RD was −0.06 (95% CLs: −0.12, −0.01), and the number needed to treat was 17 (95% CLs: 8, 100).

**Any Intracranial Hemorrhage**

Three studies reported this outcome, all of which were early rescue treatment studies. There was no evidence of an effect of iNO on overall intracranial hemorrhage frequency (typical RR: 1.0; 95% CLs: 0.73, 1.37).

**Severe Intracranial Hemorrhage**

Six of the early rescue treatment studies reported this outcome, and they showed a trend toward increased incidence of severe intracranial hemorrhage (RR: 1.27; 95% CLs: 0.99, 1.62; RD: 0.06; 95% CLs: 0.00, 0.13) (Fig 3). Because most intracranial hemorrhage...
occurs in the first 3 days of life, the studies with later entry would not be expected to demonstrate an effect on intracranial hemorrhage. Evolution of preexisting abnormalities, development of hydrocephalus, and occurrence of periventricular leukomalacia were reported as a single variable by Ballard et al., and results were not different between the groups. Of the studies of routine use for intubated preterm infants, only Kinsella et al. reported severe intracranial hemorrhage as a separate outcome, which was not affected (RR: 0.77; 95% CLs: 0.55, 1.09; RD: −0.04; 95% CLs: −0.08, 0.01).

Severe Intracranial Hemorrhage or Periventricular Leukomalacia

The early rescue treatment studies showed no significant effect, but there was a trend toward an increase in the frequency of this adverse outcome (RR: 1.16; 95% CLs: 0.93, 1.44; RD: 0.04; 95% CLs: −0.02, 0.10) (Fig 4). The studies of routine use for intubated preterm infants

![Forest plot of the effects of iNO on severe intracranial hemorrhage in preterm infants.](image)

![Forest plot of the effects of iNO on the combined outcome of severe intracranial hemorrhage or periventricular leukomalacia in preterm infants.](image)
showed a reduction in the frequency of this outcome (RR: 0.70; 95% CLs: 0.53, 0.91; RD: −0.07; 95% CLs: −0.12, −0.02; number needed to treat: 14; 95% CLs: 8, 50).

Neurodevelopmental Outcomes
To date, the only studies to report on neurodevelopmental outcomes were the studies by Mestan et al27 and Bennett et al28 and the INNOVO trial.18 From the original study by Subbedar et al,23 22 children were still alive at 30 months of age and 21 (7 iNO-treated infants and 14 control infants) were examined formally.28 There were no significant differences in outcomes. The definition of “severe neurodisability” in that report was very similar to our definition of neurodevelopmental disability; the 5 infants with severe neurodisability (Mental Developmental Index or Psychomotor Development Index of <71, cerebral palsy, or sensorineural impairment) were all control infants.

The study by Mestan et al27 showed a significant reduction, at corrected age of 2 years, in the frequency of a composite outcome of neurodevelopmental disability (cerebral palsy, bilateral blindness, bilateral hearing loss, or a score on the Bayley Scales of Infant Development >2 SDs below the mean). This improvement was largely the result of a decrease in the incidence of Bayley Scales of Infant Development scores >2 SDs below the mean. Cerebral palsy rates were not different between the groups.

The INNOVO trial investigators reported the rates of major disabilities at 1 year of age, which were not different between the groups.16 Severe disability was defined as no/minimal head control or inability to sit unsupported or no/minimal responses to visual stimuli (equivalent to a developmental quotient of <50). There was no difference between the groups (7 of 55 patients vs 3 of 53 patients).

Severe Retinopathy of Prematurity
There was no evidence of an effect on severe retinopathy of prematurity (reported only by Mestan et al27).

Retinopathy Requiring Surgery
None of the groups of studies showed an effect on this outcome. For the studies with early rescue treatment, the typical RR was 0.86 (95% CLs: 0.58, 1.29; RD: −0.02; 95% CLs: −0.07, 0.03). For the studies with entry after 3 days on the basis of BPD risk, the typical RR was 1.04 (95% CLs: 0.78, 1.38; RD: 0.01; 95% CLs: −0.06, 0.08). For the studies of routine use for intubated preterm infants, the typical RR was 1.09 (95% CLs: 0.79, 1.50; RD: 0.01; 95% CLs: −0.04, 0.06).

Sensitivity Analysis
Repeating the analyses for the early rescue treatment group after exclusion of the 2 studies that allowed backup treatment of control subjects did not affect any of the analyses substantially. There remained no significant effect on the incidence of survival without BPD (RR: 0.94; 95% CLs: 0.88, 1.02), with a similar trend toward increased incidence of grade 3 or 4 intraventricular hemorrhage or periventricular leukomalacia (RR: 1.15; 95% CLs: 0.89, 1.48).

DISCUSSION
Overall Findings
This review suggests that there may be identifiable groups of preterm infants who receive substantial benefits from iNO therapy, with reduction in the incidence of brain injuries visible on ultrasound scans and potential reduction in mortality rates; however, the precision of the estimates of these effects is low and the number needed to treat may be large, if these effects are confirmed. There are other groups of infants with evidence of adverse effects (specifically, increased rates of severe intracranial hemorrhage) without evidence of benefits.

Death
None of the individual trials showed a reduction in mortality rates. Only the meta-analysis of the 2 trials that evaluated routine use for intubated infants demonstrated a potential difference. In that case, the effect was marginally significant. The typical RR was 0.77 (95% CLs: 0.60, 0.98). The RD was −0.06 (95% CLs: −0.11, −0.01); although this is potentially an important magnitude of effect, the estimate lacks precision. The number needed to treat to save 1 infant may be as few as 9 infants or as many as 100 infants. The groups of studies of early rescue treatment and later treatment for infants at risk of BPD showed no effect on mortality rates.

Survival Without BPD
There was no apparent benefit of iNO in early rescue treatment studies. A preplanned subgroup analysis of the largest of the rescue studies, that reported by Van Meurs et al,17 suggested that there might be a reduction in the incidence of the combined outcome of death or BPD among infants with birth weights of >1 kg but not either outcome separately. The early routine treatment studies showed a modest and barely significant reduction in the incidence of the combined outcome of death or BPD. There was heterogeneity in this outcome ($^2 = 64.6\%$), with the larger study by Kinsella et al12 showing no significant benefit in their overall analysis. A preplanned subgroup analysis from that study suggested that the infants at lower risk (birth weights of >1000 g) experienced benefit in this outcome. In the study by Schreiber et al21 although the overall analysis indicated significance, the subgroup analysis suggested that it was only the less-sick infants (with OIs less than the median) who benefited. Therefore, it may be that, to reduce the
risk of BPD, therapy must be instituted before there is major lung injury.

The study by Ballard et al reported a significant benefit of iNO in improving rates of survival without BPD. However, the figures given in the article (165 of 294 vs 182 of 288 patients) were not significant when RevMan software was used to calculate the RR and its CLs, as shown above, or when SPSS software (SPSS, Chicago, IL) was used to perform a simple chi-squared test (without correction for continuity; \( \chi^2 = 2.841; P = .092 \)) or to perform unadjusted, univariate, logistic regression analysis (odds ratio: 0.75; 95% CLs: 0.54, 1.048; \( P = .092 \)). The reason for this discrepancy is not clear.

A post hoc subgroup analysis of the study by Ballard et al showed a significant reduction in the incidence of the combined outcome of death or BPD for infants who were 7 to 14 days of age at randomization; infants with less-severe disease also might be more likely to benefit (although the interaction term for that analysis was not significant). These findings suggest that additional studies, focusing on infants who seem, from these subgroup analyses, to receive the greatest benefit, may be worthwhile.

### Brain Injury

The early rescue treatment studies showed no significant effect on brain injury, as documented on ultrasound scans, but there was a trend toward increased rates of serious intracranial hemorrhage. Some of the studies, including the largest study in the group, did not require head ultrasound scans before enrollment; however, there were no major baseline group differences in that randomized trial, and the only likely explanation for the finding is that it was indeed a treatment effect. Although the result was not significant according to the conventional strict threshold of \( P < .05 \), potential evidence of harm should always be taken seriously, especially when there is no evidence of benefit. The studies with later entry on the basis of BPD risk would not be expected to demonstrate effects on intracranial hemorrhage incidence.

The studies of early routine use for intubated preterm infants showed a reduction in the incidence of serious, ultrasonographically diagnosed, brain injury (either severe intracranial hemorrhage or the combined outcome of severe hemorrhage or periventricular leukomalacia). The number needed to treat for this important outcome was a modest 14, but the confidence interval was wide (95% CLs: 8, 50). Although the criteria for study entry reported by Kinsella et al included gestational age of \( \leq 34 \) weeks, the mean gestational age was actually 25.6 weeks for each group and the mean birth weight was <800 g, demonstrating that higher-risk groups were enrolled. Similarly, in the study by Schreiber et al, the actual birth weight was \( \sim 1 \) kg for the 2 groups and the gestational age was \( < 28 \) weeks; also, that hospital serves a somewhat deprived, inner-city neighborhood with a low rate of prenatal steroid use (<60%).

Clearly, if a population at very low risk for serious brain injury were treated, then the absolute benefit of iNO would be substantially less; for example, it would be very difficult to show a reduction in the incidence of severe brain injury in mildly ill infants born at 30 to 32 weeks’ gestation. Additional analysis of the patient characteristics that predict a beneficial response is therefore important. It has been suggested that there may be ethnic differences in responses to nitric oxide, accounting for some of the different results of the studies; however, the study of Kinsella et al did not show an effect of ethnic group. Infants who are less severely sick and of higher birth weight but still at risk of adverse outcomes seem to have the greatest benefit, on the basis of the currently available data.

### Neurologic and Developmental Outcomes

Neurodevelopmental outcomes were not improved in the only early rescue treatment study to report outcomes. Outcomes have not been reported by most of the studies performed to date. For the early routine use studies, Schreiber et al demonstrated a reduction in the incidence of abnormal neurodevelopmental outcomes at 2 years of age, largely attributable to an improvement in Bayley Scales of Infant Development scores. Kinsella et al, who also showed an improvement in the ultrasound appearance of the brain, have not yet reported longer-term outcomes.

### CONCLUSIONS

Our systematic review of the literature revealed that, in the group of studies of rescue therapy of very sick preterm infants who met criteria for poor oxygenation, iNO did not improve rates of survival, survival without BPD, or brain injury. In fact there was some evidence of an increase in the frequency of severe intracranial hemorrhage and of the combined outcome of severe intracranial hemorrhage or periventricular leukomalacia. In view of these findings, iNO should not be used routinely as rescue therapy in cases of severe respiratory failure among preterm infants. Although there are variations in eligibility criteria, doses of iNO, and duration of therapy, there is no clear indication from the early rescue treatment studies that this approach to treatment is promising.

In view of the lack of statistically significant benefit and the lack of long-term follow-up data from the group of studies of the later use of iNO for infants at risk of BPD, iNO use in this clinical situation cannot be recommended presently. Additional studies restricted to infants determined to be most likely to benefit, on the basis of the subgroup analyses described above, are warranted.

In contrast, the group of studies of early routine use of
iNO for preterm infants undergoing ventilation who were not severely ill but nevertheless were at risk for serious brain injury or BPD showed promise. There was heterogeneity in the effects on BPD, with the main benefit being evident from a single-center study.\textsuperscript{21} Other studies to confirm this effect and to demonstrate its generalizability are required. Furthermore, only that study reported longer-term neurodevelopmental outcomes, and caution is suggested before more-widespread implementation of this use of iNO. Clear criteria for which patients in this population should be considered for treatment do not exist currently. The apparent decreases in the frequency of the combined outcome of death or BPD and in ultrasonographically detected brain injury suggest that infants who are at significant risk for these 2 outcomes but are not seriously ill would be the most appropriate target subjects for additional studies with long-term follow-up monitoring. Confirmation of the efficacy of such an approach would raise questions regarding why such infants would benefit but more-seriously ill infants would not. It is possible that infants who are sick enough to fulfill the entry criteria of the rescue studies have already suffered brain and pulmonary injuries that are too severe to be improved with iNO, whereas routine “prophylactic” use may be able to reduce the incidence of such injuries. This possibility warrants additional research.

REFERENCES

23. Subbedar N, Ryan S, Shaw N. Open randomised controlled trial of inhaled nitric oxide and early dexamethasone in high risk preterm infants. \textit{Arch Dis Child Fetal Neonatal Ed.} 1997;77:F185–F190
ABSTRACT

Stimulant medications (amphetamine and methylphenidate) are the best-documented treatments for attention-deficit/hyperactivity disorder, but their short pharmacokinetic and behavioral half-lives have historically produced irksome time-course effects. New drug-delivery systems designed to eliminate the need for frequent dosing include the methylphenidate transdermal system, in which the matrix acts as both the drug reservoir and the skin adhesive. The methylphenidate transdermal system patch, in contrast to long-acting oral preparations, requires a paradigmatic shift in clinical thinking, as well as refinement of clinical management skills. For dosing with the methylphenidate transdermal system patch, clinicians must think in terms of a retrievable form of drug delivery (in milligrams per hour) rather than a fixed nonretrievable dose (in milligrams per dose or milligrams per day). Clinicians and patients can determine the optimal clinical dose by controlling 2 variables: (1) patch size (controlling milligrams per hour) and (2) duration of patch wear. The new paradigm is worth learning, because the patch offers several advantages over oral preparations for some patients, chiefly individualized control over effect duration (determined by when the patch is applied in the morning and removed in the afternoon/evening). Taking full advantage of this treatment option requires educating the patient and parents regarding practical elements of daily use. These elements include patch-site selection, application techniques, management of wear time to optimize the daily time course of clinical benefits, and skin hygiene. This article summarizes clinical principles that physicians may find useful in managing this new addition to the attention-deficit/hyperactivity disorder treatment armamentarium.
SYMPATHOMIMETIC STIMULANT MEDICATIONS (ie, various forms of methylphenidate and amphetamine) have long been recognized as the most effective and best-documented treatment for attention-deficit/hyperactivity disorder (ADHD). ADHD is a neurodevelopmental syndrome of chronic inattention, distractibility, impulsiveness, and restless overactivity, which impairs functioning at home, in school, and in the community (and often later in life on the job or in a marriage). Stimulants have been the medication of choice for the ADHD syndrome since 1937, when Bradley reported conduct and academic improvement with stimulant use in children with behavioral and emotional disturbances. Expert consensus guidelines continue to support the use of stimulant medications as first-line therapy for ADHD. The risk/benefit ratio is highly favorable and well documented, as evidenced in a 1996 meta-analysis that evaluated 155 controlled studies of 5768 children, adolescents, and adults; Spencer et al. found an average efficacy of 70% for stimulant medications. Goldman et al. on behalf of the Council on Scientific Affairs for the American Medical Association, concluded that, with appropriate dose titration, up to 90% of children properly diagnosed as having ADHD respond to at least one stimulant without experiencing a major adverse event.

Unfortunately, stimulants as originally manufactured (ie, immediate-release forms) have a short half-life and limited duration of effect, which make the management of medication time-course effects somewhat challenging. The need to give a noon dose at school made the immediate-release forms of the stimulants both onerous and stigmatizing for pediatric patients, parents, and schools. Furthermore, it has invited public criticism and political fire from some quarters. The need to remember the exact timing of repeated doses is especially problematic and impractical for a disorder for which forgetfulness in daily activities is 1 of the 18 defining symptoms.

In response to these problems, a number of longer-acting preparations and delivery systems have been developed to level the diurnal curve and to eliminate the need for school-time dosing. These developments include coated or other forms of sustained-release tablets or pills, encapsulated beads with time-release coatings, an oral osmotic delivery system, a new longer-acting molecule (atomoxetine), a conjugated d-amphetamine molecule (lisdexamfetamine), and a transdermal patch (methylphenidate transdermal system [MTS]).

Use of the new, long-acting, oral, stimulant preparations requires only minor shifts in thinking by clinicians (essentially aggregation of short-acting doses administered 2 or 3 times per day into a single, long-acting dose administered in the morning). However, management of ADHD by using the MTS patch involves a paradigmatic shift in thinking, from a fixed, nonretrievable, oral dosage administered in milligrams per day to a flexible, removable, transdermal dosage form prescribed in terms of hourly drug-delivery rate (ie, milligrams per hour) and adjustable patch wear time (total milligrams per day). The benefits of this paradigm shift are increased flexibility and control over the time course of clinical effects and the duration of benefits and adverse effects, possible avoidance of some gastrointestinal problems, and more-uniform systemic drug delivery than with most oral treatments. As is characteristic of transdermal drug delivery, compared with oral dosing, MTS patches have more consistent absorption and systemic bioavailability over the wear period than do most oral preparations. Table 1 lists some of the advantages and disadvantages of the MTS patch relative to long-acting oral preparations.

One advantage of the MTS patch is its possible lower abuse potential. Some parents fear that use of stimulant medications may encourage their children to abuse other drugs. Clinicians can assure parents that there is no published evidence that stimulant therapy for ADHD in childhood is associated with increased risk of subsequent substance use disorders. Claims that early stimulant treatment might actually be protective against later abuse have been questioned. Most long-acting oral stimulants some-

<table>
<thead>
<tr>
<th>TABLE 1 Advantages and Disadvantages of the MTS Patch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td>Retrievable dosage form that allows individualized patient and clinician control over onset and duration of effects and side effects</td>
</tr>
<tr>
<td>All-day coverage lasting through evening (can have longer duration than current oral methylphenidate preparations)</td>
</tr>
<tr>
<td>Absorption curve smoother than that of oral medication boluses</td>
</tr>
<tr>
<td>Absorption completely unaffected by meals or competitive metabolism of coadministered medications</td>
</tr>
<tr>
<td>No need to swallow tablets or capsules</td>
</tr>
<tr>
<td>Possible application by parent before child waking, to optimize early-morning functioning</td>
</tr>
<tr>
<td>Probably more difficult to divert or to abuse than oral preparations</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Although long-acting oral preparations of methylphenidate offer only dosage manipulation as a method of titration, the MTS patch offers manipulation of both hourly drug-delivery rate and wear time as possible means of titration for optimal clinical exposure. The patch “dose” has 2 independent components, namely, the size of the patch, from 12.5 cm² (nominally 10 mg) to 37.5 cm² (nominally 30 mg), and adjustable wear time, which together determine the milligrams of methylphenidate absorbed per day by a given patient. Naturally, interpatient variations in absorption rates also affect the effective dose and must be considered whether skin or gut mucosa is the absorption route. Skin absorption differences are minimized by consistent selection of proximal application sites (hip) for each patient. Distal sites (arm or leg) have more absorption variability both between and within patients.

As with oral preparations of methylphenidate, clinicians should begin treatment of stimulant-naive pediatric patients, especially those with small body mass, with the lowest convenient amount, the 12.5-cm² MTS patch. A wear time of 9 hours with this size provides a nominal total daily dose of ~10 mg, but this has the same effect as 15 mg (5 mg administered 3 times per day) administered orally because it avoids first-pass liver metabolism. This is approximately equal to an 18-mg osmotic delivery capsule, which is not absorbed completely. Although it is not recommended in the package insert, treatment can be initiated with larger patches for patients changing from oral doses of >20 mg/day. For example, the 18.75-cm² patch delivers a 15-mg nominal dose in 9 hours of wear, which has the same effect as a 27-mg osmotic capsule or a 20-mg encapsulated bead preparation. Table 2 lists the doses of the MTS patch that have been investigated and their approximate oral immediate-release methylphenidate equivalents.

A 9-hour wear time should be tried initially (Fig 1), with increases or decreases according to how rapidly the effect wears off after patch removal (usually ~3 hours). Generally, patch removal 3 hours before bedtime results in an appropriate wear time (usually ~9–11 hours for schoolchildren), although the patch can be worn for up to 16 hours if extended effects are needed. It is useful to remove the patch well before bedtime, to allow a good appetite for a bedtime snack. The clinical effects are assessed within 1 week and, if the results are insufficient

<table>
<thead>
<tr>
<th>MTS Patch Size, cm²</th>
<th>Drug-Delivery Rate, mg/hᵃ</th>
<th>MTS Apparent Dose With 9-h Wear Time, mg</th>
<th>IR Methylphenidate Equivalent, mg tidᵇ</th>
<th>Methylphenidate Osmotic Release, mg</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.50</td>
<td>1.10</td>
<td>10</td>
<td>5.0</td>
<td>18</td>
</tr>
<tr>
<td>18.75</td>
<td>1.65</td>
<td>15</td>
<td>7.5</td>
<td>27</td>
</tr>
<tr>
<td>25.00</td>
<td>2.20</td>
<td>20</td>
<td>10.0</td>
<td>36</td>
</tr>
<tr>
<td>37.50</td>
<td>3.30</td>
<td>30</td>
<td>15.0</td>
<td>54</td>
</tr>
</tbody>
</table>

IR indicates immediate-release; tid, 3 times per day.
ᵃIn vivo delivery rate for pediatric patients when the patch is applied to the hip.
ᵇBecause some of the orally administered immediate-release methylphenidate is eliminated in the first pass through the liver, 15 mg administered orally over a 12-hour period does not equal 15 mg absorbed from the MTS patch.
during the morning and afternoon, then the patch size is increased. If the clinical benefits are satisfactory during the school day but not in the evening, then the evening can be covered merely by increasing the wear time without changing the patch size. It must be remembered that increasing the wear time without changing the patch size increases the daily dose.

Systematic monitoring of clinical and adverse effects remains an essential component of treatment and facilitates dosage titration. The American Academy of Pediatrics, in conjunction with the National Initiative for Children’s Healthcare Quality, developed a toolkit that facilitates diagnosis of ADHD and systematic monitoring of therapy.15

If the clinical benefits are satisfactory but flawed by interference with sleep or other late-day adverse effects, then the wear time can be decreased gradually to whatever time allows for cessation of the adverse effects. It should be noted, however, that the adverse effects of insomnia and anorexia, which are noted with all stimulants at the initiation of treatment, have been reported to resolve with repeated use, even without changes in wear time, in 40% to 60% of cases.16,17 This finding is consistent with the results for oral dosing. Some children may have such slow metabolism that a removal time of 3:00 to 4:00 PM, with subsequent drug washout, may provide adequate duration of effect. At each weekly

---

**FIGURE 1**
Clinical management guidelines for the MTS patch.
assessment, decisions should be made regarding both patch size (milligrams per hour) and wear time (total milligrams per day), on the basis of the individual patient’s response, until the optimal patch size and duration of wear are determined.

If an optimal size is not found among the commercially available sizes, then the question of whether an “in-between” size could be home-manufactured by snipping a little off a larger size naturally arises. The protocols in the premarketing studies did not allow for such a strategy; consequently, there has been no evidence to support it. However, there is no logical reason why this could not be done, because the methylphenidate is dispersed uniformly throughout the adhesive. If this is done, then the corners should be rounded, to prevent catching on clothes. Once the correct patch size and wear time are determined for the individual, the patient should be assessed (with inspection of the patch site) every other week for 1 month and monthly thereafter.

The beginning and the end of the day deserve special consideration. Like extended-release oral medications, transdermal delivery of methylphenidate (particularly starting doses) may have a somewhat slower onset of clinical effect than an immediate-release (immediate-release oral preparations. To compensate for the slower onset, the patch should be applied a bit earlier in the morning than an immediate-release oral preparation would be administered. It is important to note that the patch tends not to interfere with breakfast appetite if application precedes breakfast by <1 hour. For a child who exhibits significant difficulty in the morning before medication, parents may choose to apply the patch before the child awakens, to allow some effect to be present on arising.

**PATCH-SITE SELECTION AND PATCH APPLICATION**

The MTS patch is designed to be worn on the lateral hip beneath the underwear, avoiding the waistline. The site and application process are illustrated in the package insert. Placement of the MTS patch on the lateral hip serves the dual purpose of hiding the patch and standardizing the expected absorption rate. It is important to note that absorption occurs at different rates in different body regions and the variability among and within patients is greater in distal areas, compared with proximal areas. Optimal absorption depends on the maintenance of the hygiene and integrity of the skin at the patch-application site and the use of proper placement techniques.

It is important that the patch be placed where tights-fitting clothing, especially belts, will not rub it. Such abrasion may cause the MTS patch to rub off or may roll up the patch edges and interfere with absorption, as well as causing irritation. The patch should be applied to clean, dry, intact skin that is free of powder, oil, lotion, cuts, abrasions, or irritation. Ordinary bathing is adequate to satisfy the “clean” requirement; the patch site should not be scrubbed or cleaned with alcohol or other agents that might compromise skin integrity. A residual adhesive outline of the patch (like a “bathtub ring” or “Band-Aid ring”) is not harmful to leave but, if parents wish to remove it for cosmetic reasons, then removal can be performed with gentle swabbing with a cotton ball with vegetable/mineral oil or lotion.

To allow skin recovery between applications, patch-application sites and hips should be alternated daily. Unless the MTS patch is too large or the hip is too small, each hip should be divided into 2 application sites (anterolateral and posterolateral) to yield a total of 4 application sites, which allows 3 days between applications at a given site. The application site rotation might be as follows: right anterior, left anterior, right posterior, left posterior, and then back to right anterior. A good rule is to allow a fingers’ breadth between the old and current application sites. Occasionally, the child is large enough and the optimal MTS patch size is small enough that each hip can be divided into 4 quadrants, which allows each patch-application site to be used only once every 8 days, but this degree of complexity would be needed only in rare cases of very sensitive skin. Any area subject to stretching or wrinkling, such as flexor/extensor surfaces (eg, inguinal area or buttocks), should be avoided, to prevent discomfort and failure of adhesion.

Proper handling of the MTS patch immediately before application can promote good patch adhesion. The patches are supplied in a sealed tray and are individually sealed in pouches. They should be cool before application, to facilitate removal of the rigid protective liner. (Ten to 15 minutes in the refrigerator may help in hot weather, although the patches are intended for storage at room temperature.) The pouch should be cut open carefully, without cutting into the MTS patch, and the patch should be applied immediately after removal from the pouch. Care should be taken to avoid touching the adhesive surface with the fingers. One half of the rigid protective liner, which is presplit diagonally into 2 sections, can be removed without touching the exposed adhesive directly by first flexing the patch 3 or 4 times to loosen the acute corner at the split. If this is not enough to loosen the acute corner, then the patch may be folded and 1 cover section pushed with a sliding motion. Peel off one half of the protective liner while holding the patch by the other half. With the intact half of the protective liner as a handle, the exposed adhesive surface of the MTS patch should be applied to the selected area of the hip, pressed firmly in place, and smoothed down. While the half of the patch already attached to the skin is held in place, the patch should be folded over. This action should separate the acute corner of the other half of the protective liner, which then can be pulled off gently. The rest of the patch should be pressed firmly in place and the entire patch then smoothed down. After application, the patch should be pressed firmly and warmed with the palm of the hand for 30 to 60 seconds, to enhance adhesion. Several studies demonstrated good ad-
herence of properly applied MTS patches in highly active, summer camp settings (eg, swimming and sports).17,18

SKIN HYGIENE

Methylphenidate, which is a sympathomimetic stimulant, is a mild irritant and dilates blood vessels, inducing reddening ("blush") of the skin under the patch, which is "normal" and clinically benign. This normal erythema should not be confused with a rash or allergic reaction. Clues to identifying normal erythema are that the redness is confined to the patch area, it is flat (with no edema or papules), and it usually resolves 24 to 36 hours after patch removal. Parents should be advised that, if erythema persists for >2 days or is associated with papules or edema, it should receive clinical attention. Figure 2 (from a dermatologic skin test, not an actual MTS patch application) illustrates the kind of rare reaction that should result in the termination of patch treatment and a switch to oral medication.

Although stimulants are not commonly known allergens, it is possible to have an allergic reaction to the adhesive (rare). Furthermore, other topical dermatitides or mechanical irritation may result infrequently from repeated patch application and removal. If the effect is severe enough, treatment with the patch should be discontinued. It is important to note that most rashes (dermatitides) and other reactions are preventable by adhering to the application instructions and paying attention to maintaining skin integrity.

In addition to rotating patch sites by alternating hips and, if possible, specific areas of each hip, several other measures can be taken to preserve skin integrity. Parents and children should be instructed regarding gentle bathing practices with hydrating lotion soaps (such as CeraVe [Coria Laboratories Ltd, Fort Worth, TX] and Cetaphil [Galderma Laboratories, Fort Worth, TX]) and the use of vegetable or mineral oil to wipe away the glue resins gently. Including children in this discussion introduces a ritual for removing the patch, akin to brushing one’s teeth. Clinicians should recommend no rubbing of the area after patch removal and resistance of the natural tendency to scrub off any residual adhesive around the patch outline.

Frictional abrasion usually irritates the skin more than does residual adhesive. Emollients such as petrolatum can also be used to hydrate the skin and to help it heal.

The recommendation for the patch site to be clean and dry does not imply the use of special cleaning agents that might be harsh on the skin; ordinary bathing with a fragrance-free moisturizing soap is adequate. If irritation of the skin develops, a simple water-soluble moisturizer or 1% hydrocortisone ointment (available without a prescription) may be applied to the most-recent patch site. The emollient/ointment should be applied only when the MTS patch is removed, allowing 4 daily showers or baths between applications and reuse of the formerly irritated site (assuming that the 4-site rotation described above is used). It is important that emollients/ointments not be used on a site just before patch application, because of potential interference with both adhesion and absorption.

PATIENT AND PARENT EDUCATION

Not only the clinician but also the patient (or, in the case of youngsters, the parent) must be aware of most of the principles described above. Table 3 provides a checklist of important information that must be communicated to patients and parents to obtain optimal results and to promote cooperation and compliance. It may be photocopied and given to them to look at while they receive instructions and then to keep as a reminder.

The child should be trained gradually to take responsibility for skin hygiene and patch application. Self-application would be performed best with the aid of a mirror, to confirm that the application site from 2 days earlier is not overlapped. As with any ADHD medication, initial visits need to be frequent and long enough not only to titrate dosing accurately but also to educate and to promote communication among all who can help the child (parent, teacher, coach, and child). A daily report card2 is a good way to promote communication between home and school. As the "star of the therapeutic team,"19 the child should take gradually more responsibility for communication, as well as monitoring of adverse effects and skin hygiene, starting with removal of the patch at the proper time with parent prompting.
CONCLUSIONS

Optimization of clinical benefits with any ADHD medication requires the ability of clinicians and patients to control precisely both the robustness of the response and the daily time course of that response. Presently, pharmacologic treatment of ADHD does not consist of a single ideal treatment, and additional treatment options are clinically valuable. The availability of the MTS patch expands clinician and patient options in several ways; it bypasses the gut and its vagaries, it eliminates the need to swallow, it allows for consistent clinical effects throughout the day for as long as needed, it produces greater duration of effect than do existing oral methylphenidate regimens, and, perhaps most importantly, it offers complete control over both onset and offset of effects, empowering the parent and the child to take control of the symptoms and promoting cooperative relationships among the child, the parent, and the clinician. The patch may be an obvious choice for patients who have trouble swallowing pills. It also should be useful for patients who require varying durations of effects from day to day, have rapid metabolism (requiring a second dose even with extended-release oral preparations), require flexible management of late-day adverse effects, or routinely require extended symptom control of longer duration than available with the oral preparations. Therefore, the time needed to master the new treatment option by learning the knowledge and skills summarized here would seem to be well invested.

REFERENCES


17. Burrows-MacLean L, Pelham WE Jr, Hoffman MT. Improved nightly parent ratings with once-daily transdermal methylphenidate. Presented at: 110th annual meeting of the American Psychiatric Association; August 22–25, 2002; Chicago, IL


The emergence and rapid spread of walk-in clinics staffed with nurse clinicians in retail stores and shopping malls challenges both medical and nonmedical communities to consider the relative priorities of continuity and the medical home compared with convenience and cost. Continuity of care has 3 dimensions: time, accessibility, and setting. The time dimension refers to having the same pediatrician (or clinician team) care for a child over a considerable period of time, optimally for many years, so that the pediatrician can develop a long-lasting relationship with the child and his or her family. Having continuous care allows for deeper understanding of the past as well as present issues that impact the child’s development, health, and well-being. The accessibility dimension refers to ensuring that the family will be able to contact the pediatrician or a member of the coverage team 24 hours a day, 7 days a week, and be able to get needed care promptly in an appropriate medical setting. The setting dimension refers to coordinating care across multiple settings including the hospital, physician outpatient office, school, home health care agency, and, if necessary, hospice program. Continuity of care is important for both pediatric primary and subspecialty care. In several situations a subspecialty pediatrician, such as an oncologist, may provide continuous care and assume the care-coordination responsibilities.

For more than 3 decades pediatric residency training programs have attempted to “teach continuity” through the establishment of continuity clinics in which residents care for a panel of patients. Because most residents only attend these clinics 1 or 2 half-days per week, continuity over time tends to be limited to scheduled preventive care and chronic care visits. However, even this limited amount of continuity is being eroded when a higher priority is given to inpatient rotations on ICUs, night-call coverage schedules, and limitations on resident working hours.

The erosion of the continuity experience in residency training is also mirrored in private practice because of several internal practice reasons and external health care system reasons. More pediatricians are practicing in larger groups and share patients among several partners. More pediatricians are no longer involved in caring for their patients who require hospitalization. More pediatricians use after-hours call centers and no longer take calls or see patients in their office or an emergency department after hours. Pediatricians currently have a more limited amount of time to spend with families during their visits. On average, they are spending only 18.3 minutes with families during a well-child care visit.1 These shorter, less comprehensive visits are less satisfying to the families and pediatricians and undermine the value of continuity. Health plans attempt to purchase services at the lowest price possible, which forces pediatricians to accept unreasonably low payments for services, and often reimburse pediatricians less money than it costs to purchase and administer vaccines. To remain financially solvent despite these low payments, pediatricians see increasing numbers of children for shorter visit times. As patient volume increases, the practice becomes less efficient, because pediatricians must to hire more
administrative staff for billing and authorization paperwork.

There are also structural problems in health care that undermine the ability to provide continuity of care. More families are being forced to change health plans yearly when they decide to change jobs or their employers choose a less costly health plan. Changing health plans often means changing physicians. Having so many children with gaps in insurance coverage also impacts continuity of care. On the basis of data from the Medical Expenditure Panel Survey, the American Academy of Pediatrics (AAP) found that 21 million children 0 to 21 years of age were uninsured for all or part of 2004; 9 million were without coverage for the entire year, and 12 million experienced gaps in coverage sometime during the year.

Although there is considerable support for the medical home concept by state and national policy makers and medical home language is being included in several state and national legislative proposals, there is a disconnect between what is happening in the real world of pediatric practice and the world that the AAP and other professional societies would like to see. Continuity is considered to be a core value of practicing pediatrics by the AAP and is the foundation of the academy’s conceptual framework of a “medical home.” What started as a concept for children with special health care needs in 1967 has been extended to all children and, more recently, to adults. In February 2007 the primary care societies, including the AAP, American Academy of Family Physicians, American College of Physicians, and the American Osteopathic Association, released a statement: “Joint Principles of the Patient-Centered Medical Home.” These principles include the following: First, each patient should have an ongoing relationship with a personal physician who has been trained to provide first-contact, continuous, and comprehensive care. Second, the personal physician should lead a team of individuals at the practice level who collectively take responsibility for the ongoing care of patients. Third, the personal physician should be responsible for providing for all the patient’s health care needs or taking responsibility for appropriately arranging care with other qualified professionals, including care for all stages of life: acute care, chronic care, preventive services, and end-of-life care. Fourth, care should be coordinated and/or integrated across all elements of the complex health care system (eg, subspecialty care, hospitals, home health agencies, nursing homes) and the patient’s community (eg, family, public and private community-based services). Fifth, care should be facilitated by registries, information technology, health information exchange, and other means to ensure that patients get the indicated care when and where they need and want in a culturally and linguistically appropriate manner.

In the real world, we are beginning to see retail-based clinics that serve adults and children open up in department stores, discount warehouses, pharmacies, and grocery stores throughout the country and are becoming popular with families. Although professional organizations and politicians embrace continuity and the medical home concepts, the emergence and expansion of these retail-based clinics demonstrate that enhancing convenience and reducing costs may have a higher priority for many families, as well as health plans, than enhancing continuity.

Supporters of these clinics say that they are a cheaper and better alternative to hospital emergency departments, especially for uninsured patients and Medicaid patients who cannot find a regular source of primary care. They also point out that it is easier for working mothers and fathers (especially single parents) to bring their children to a retail-based clinic after work where they can also do some needed shopping. One-stop shopping and health care may result in some children receiving care that they otherwise might not get. Most of the clinics say they have referral or consultative arrangements with local physicians or clinics. Detractors voice concerns about whether the consultative and referral arrangements will be adequate and that the clinics will fragment care and disrupt attempts to have a unified medical chart. Will the nurses who staff these clinics have the training to care for children with special health care needs who present with an acute illness or to correctly diagnose more severe conditions that present with common symptoms? Concerns have also been raised that nurses who staff the clinics in drug stores will be encouraged to prescribe antibiotics and other medications. Drug stores may be willing to subsidize the operation of these clinics to capture the income from the sale of pharmaceuticals.

Pediatricians must face the challenge of retail-based clinics to continuity and the medical home by enhancing continuity in pediatric training programs and pediatric practice and by demonstrating how continuity improves quality. Both training programs and practice must strive to achieve continuity and improve access to care, quality of care, practice efficiency, convenience, and patient satisfaction. The “Joint Principles of the Patient-Centered Medical Home” statement and the AAP statement on retail-based clinics both encourage physicians to enhance access with open scheduling, expanded hours (evenings and weekends), and new options for communication among patients, their personal physician, and practice staff. The emergence and expansion of well-funded retail-based clinics requires a reappraisal of how we practice and how public health programs and private health plans pay for services. Our challenge is to clearly demonstrate the value of continuity and the medical home on quality and cost. Failure to meet this challenge will allow retail-based clinics to expand and become a dominant player in the marketplace.
REFERENCES


THE NEONATAL community deserves congratulations for responding vigorously to Silverman’s1 call for randomized controlled trials (RCTs) to evaluate neonatal therapies. Although more trials are still needed,2 existing RCTs present new challenges in interpretation. One of the most vexing is when to proclaim innovative therapies as “standard of care.”

The neonatal critical care community faces this challenge in evaluation of hypothermia as treatment for hypoxemic-ischemic encephalopathy (HIE).3–5 National bodies have made declarations that the neonatal community should consider hypothermia experimental pending completion of current ongoing trials.6–9 Although the influence of these bodies is considerable, individual physicians and sites apparently feel pressure to “do something” in the very dire circumstances of HIE in the newborn. In an informal sample of convenience, we have found that some centers are performing cooling, either with or without informed consent. Although many clinicians concur with the leading bodies that state there is a need for additional trials, it is confusing for practicing neonatologists when some members of these bodies also publicly state that they are actively providing cooling therapy.

If leading centers are promoting active cooling, they have, in effect, adopted cooling as a standard of care. This may not only have legal implications but also raises ethical issues for those who believe the right thing to do currently is to continue performing RCTs. The countervailing argument is that to not offer cooling as standard therapy for such a devastating disease as HIE is itself, unethical. These opposing viewpoints are not easily resolvable except by considering what the overall benefit of eliminating residual doubt, one way or the other, would be. Our concern is that advocacy of hypothermia as a standard of care represents an excessively low threshold for accepting promising therapies and will ultimately lead to resources devoted to useless interventions that should be devoted to developing and implementing useful ones.

We argue for a conservative approach to declaring a therapy to be standard of care. Although individual clinicians may choose to implement a therapy for which the magnitude of benefit remains uncertain, designating a therapy as a standard for quality care mandates the presence of very strong evidence. In this commentary, we explain the reasoning behind our approach and propose guidelines for considering when a body of meta-analytical data is strong enough to argue that it is unethical to further randomize into ongoing, or new trials.

Why do we advocate a conservative stance? The history of physicians’ repeated endorsement of therapies that later proved useless or harmful,10 including therapies that seemed promising in RCTs,11 provides one compelling rationale. Ioannidis11 reported, after reviewing
highly cited publications of efficacious studies, that 32% were later found to have been contradicted or to have had initially stronger effects. The studies that were highly cited and not refuted had a median sample size of 1542, as opposed to those that were either contradicted or claimed initially stronger effects, which had a median sample size of 624.

In addition to the sobering lessons of subsequent reversal of initially promising results, our arguments for caution regarding hypothermia rest on limitations in the internal validity of the 2 pivotal trials, the Gluckman et al Cool Cap study,1 and the Shankaran et al National Institute of Child Health and Human Development (NICHD) whole-body–cooling study,4 and 1 smaller pilot study of 65 infants by Eicher et al.5 We are aware of 1 other pilot study12 and 1 completed but unpublished study.13 A final study, the Trial of Whole Body Hypothermia for Perinatal Asphyxia (TOBY),14 has completed recruitment but still has to achieve target end points of 18 months’ outcome. We consider the implications of results of these smaller and unpublished studies in “Have the Pooled Studies Achieved an Optimal Information Size?” below.

Researchers in the Cool Cap study found, with a subgroup analysis, a reduction in 18 months’ adverse outcome. Using whole-body cooling, Shankaran et al found an overall reduction in adverse outcomes at 18 months (relative risk [RR]: 0.72 [95% confidence interval (CI): 0.54–0.95]; P = .01). Although similar in basic goals, the 2 studies had important differences6–7 in how they achieved cooling (whether head or whole-body4,5 cooling) and in the entry criteria. The Cool Cap trial used amplitude-integrated electroencephalography (aEEG) to discern whether infants were affected enough to randomly assign them.1 Although these differences do not necessarily preclude a pooled analysis of all enrolled infants irrespective of subgroup, making a total of 478 infants; infants cooled had a 0.76 RR (as compared with control infants) for the outcome of death or disability at 12 to 18 months.15

Although these results seem compelling on the surface, 4 key concerns remain: (1) the potential for biases that arise within an unblinded study; (2) concern about the management of control-group patients; (3) how to interpret subgroup analysis; and (4) the relatively small number of patients studied to date.

1. Potential for biases arising: Both trials used a composite outcome of death and/or significant (severe in the Cool Cap trial and moderate or severe in the NICHD trial) sequelae in survivors at 18 months. Blinding in the NICU was impossible to undertake for practical and ethical reasons.

Any unblinded trial risks bias in cointervention and the process of establishing outcome events. In this case, the concern is particularly serious and arises from the question, “How do infants with severe HIE die in the NICU?” A frequent mode of death in this setting is a parental or clinician decision to withdraw care. Thus, there is a possibility that whether infants survive is a decision that, to some extent, is in the hands of the unblinded attending physicians.

Defining neurologic criteria for withdrawal of support is difficult, because there is no agreed-on definition for brain death in neonates. Moreover, the process is emotionally traumatic for all concerned. Therefore, one might question the possibility of limiting bias and increasing the transparency of the decision. In fact, in a research context, investigators could build in an independent arms-length review after a withdrawal of care by using review of the medical charts. Investigators could take the crucial step of blinding this adjudication, which could characterize the decision in terms of the certainty of a poor prognosis and the involvement of the clinician in the decision. Blinded adjudicated outcome would go a long way toward resolving concerns about differential application of criteria in intervention and control groups.

The bias we are suggesting might lead to an increase in severe morbidity in survivors, which is a result that did not occur. Nevertheless, failure to observe increases in disability in intervention groups does not exclude the possibility of underlying bias.

Two major ongoing randomized trials, the TOBY14 and Infant Cooling Evaluation (ICE),16 also lack a priori criteria for withdrawal of support or blinded adjudication of withdrawal decisions. Thus, even after those trials are complete, the issue of possible bias in withdrawal of support will remain. Therefore, it will be crucial that these trials report the incidence of death as a result of withdrawal in the intervention and control groups.

2. Were control patients optimally managed? Hyperthermia after a cerebral insult is associated with worse outcomes.17–19 Although attention has focused on the effect of cooling in the experimental arm, there is a potential that scrupulous attention to ensuring that the infant does not get overheated, rather than cooling, might be the true mechanism of benefit. In the Shankaran et al trial,4 in 41 (39%) of the 106 control infants there was at least 1 esophageal temperature that exceeded 38°C. The network investigators recently presented further analysis of core- and skin-temperature data from their control group and indicated that the range of median core temperatures in their controls was 36.3 to 38.9°C.

Furthermore, an increase in only 1°C in peak core temperature was associated with a fourfold increase in death or disability.19 Again, we were not told the exact frequency of this potential confounder in the
Cool Cap study controls, but the report suggested that it may have been as high as 23%. Avoiding hyperthermia may be challenging. Current skin temperature–based servoregulation may not be suited to the task of avoiding core hyperthermia. Some active cooling mechanism may be required to avoid excessive temperatures. In addition, the extent to which infants tolerate an upward deflection of temperature is unknown. These issues remain to be ones that the neonatal community needs to actively investigate. One could argue that avoiding hyperthermia is a novel intervention with even less evidence than cooling. On the other hand, one might view avoiding hyperthermia as a standard of care, as reflected in Neonatal Resuscitation Program guidelines. The Cool Cap investigators themselves pointed out that the interaction between severity of aEEG changes and treatment outcome shows a P value of .075, which is above the conventional threshold for significance. Thus, the apparent subgroup effect may represent a “siren song” that is best ignored. Certainly, we cannot consider it established.

3. Interpretation of subgroup analysis: Methodologists have been aware for more than 15 years of the dangers associated with subgroup analysis. In the Cool Cap study, the primary outcome of death and/or disability at 18 months did not reach the conventional threshold of statistical significance (55% vs 66%; P = .10). However, a subgroup analysis performed according to severity at presentation showed a reduction of adverse events in the moderately affected group (48% vs 66%; P = .02), but this was not the case in the subgroup of infants who were severely affected at entry (79% vs 68%; P = .51).

Under what circumstances can we be confident in the findings of a subgroup analysis? Of 7 criteria for judging the credibility of a subgroup analysis, the Cool Cap trial fails to meet 2 key criteria. First, the subgroup difference was not consistent across studies. Shankaran et al did not use aEEG at entry; nonetheless, a comparison by severity according to clinical assessment enabled an analysis (moderate HIE RR: 0.69 [95% CI: 0.44–1.07], P = .09, and severe HIE RR: 0.85 [95% CI: 0.64–1.13], P = .24). Thus, although the apparent effect was slightly greater in the moderate than in the severe group in the Shankaran et al study, the difference was small and does not substantiate the clear difference in effect claimed by the Cool Cap trial researchers. Although they enrolled exactly the same population and classifying patients according to the aEEG might have led to replication of findings, confidence in subgroup effects requires replication, which is not currently available.

Second, one can conduct a statistical analysis to determine if the difference in subgroups is compatible with the play of chance. From an independent analysis, the US Food and Drug Administration reported that “no conclusions could be drawn from the sponsor’s pooled subpopulation because the overall treatment-by-interaction test was not statistically significant.” The Cool Cap investigators themselves pointed out that the interaction between severity of aEEG changes and treatment outcome shows a P value of .075, which is above the conventional threshold for significance. Thus, the apparent subgroup effect may represent a “siren song” that is best ignored. Certainly, we cannot consider it established.

4. Have the pooled studies achieved an optimal information size? Up to now, methodologists and systematic reviewers have given limited thought to the issue of a threshold for when enough data have accumulated to conclude that a question has been answered adequately. A number of authors have highlighted the dangers of overestimating treatment effects in individual randomized trials that are stopped early, after an interim analysis. To guard against this, it is common to see trial organizers using formal stopping boundaries such as the so-called Lan-DeMets α spending function rule. Formal stopping rules represent one response to awareness that repeated looks at data from RCTs violate the fundamental assumptions that underlie conventional statistical analysis, which invalidates the conventional rule of significance (P < .05) and makes the likelihood of a false-positive finding and overestimation of treatment effects extremely high.

Systematic reviews and meta-analysis run the risk of a very similar phenomenon. Nowadays, thousands of randomized trials are conducted each year. Inevitably, some RCTs, particularly if their sample size is relatively small and they accrue relatively few events, will demonstrate spurious overestimates of effect. In a smaller but still-appreciable number, the first several small trials will produce spurious overestimates of benefit. Thus, meta-analyses represent a parallel situation of accumulating data in which early apparent benefits that come from relatively small numbers of patients may represent misleading chance phenomena.

How can one guard against these false-positive conclusions? Pogue and Yusuf first proposed a meta-analytic approach analogous to stopping rules for individual trials. Building on this early work, subsequent investigations have suggested a calculation of an optimal information size to estimate the extent of this risk of overestimates of treatment effect that arise from small data sets. These approaches remain underutilized, particularly in the neonatal community, which has failed to give adequate attention to this issue.
To perform a calculation of the “optimal information size,” one needs to know the control event rate. Pooling the control rate events of the Cool Cap, NICHD, and Eicher et al trials estimates a rate of 61.3%. Because treatment effects are rarely higher than 25% in medicine, one can assume a plausible RR reduction of 20% for death and disability with cooling. Assuming such a plausible 20% RR reduction, an α error of .05, a β error of 10%, and a control event rate of 61.3%, the optimal information size in this case would include studies of a total of 692 patients. Using a sensitivity-analysis approach, if the event rate was lower, say at 50%, the optimal information size would be 1102. Both of these estimates are greater than the 442 patients included in the 2 fully published highest standard relevant RCTs being considered. If the Eicher et al trial is included, the total recruited comes to 507. Currently unpublished trials include the TOBY, which enrolled 325 infants who are awaiting outcome at follow-up,14 and the Shao et al13 trial of 178, which had an unbalanced randomization with 111 cooled infants versus 67 control infants. Finally, 1 small pilot randomized trial12 enrolled 22 infants. Even ignoring the concerning issues of potential bias we have highlighted, additional reports and studies are required to provide a robust assessment of the effect of cooling. This is one reason to welcome the timely completion of the ICE trial with a sample size of 276 infants.16 In addition, the ICE will provide information on the feasibility and safety of a pragmatic approach to whole-body cooling in transport.

In summary, exciting potential exists in hypothermia for cooling. Is the evidence sufficiently strong that clinicians impressed with the results may cautiously use this treatment for neonatal encephalopathy while they wait for the many questions around its optimal use to be answered? Certainly. On the other hand, the neonatal community continuing with a conservative approach to declaring a new standard of care will avoid unfortunate mistakes of premature dissemination of experimental management strategies. In both adults and children with traumatic brain injury, cooling has not fulfilled its earlier promise. We should demand strong evidence of robust, consistent effects in highly valid studies that have enrolled adequate numbers of patients before mandating a new therapy for management of all relevant patients. The evidence for cooling fails to meet this standard.

ACKNOWLEDGMENTS

Dr Barks, as a site investigator for Mott Hospital, recruited infants for the Cool Cap study; Dr Kirpalani was a site investigator for McMaster University Hospital and recruited infants for the ICE study.

We acknowledge valuable methodologic advice from Dr Barbara Schmidt and Dr Edmund Hey.

REFERENCES

8. American Heart Association. 2005 American Heart Association (AHA) guidelines for cardiopulmonary resuscitation (CPR) and emergency cardiovascular care (ECC) of pediatric and neonatal patients: pediatric basic life support. Pediatrics. 2006;117(5). Available at: www.pediatrics.org/cgi/content/full/117/5/e6989
9. International Liaison Committee on Resuscitation. The International Liaison Committee on Resuscitation (ILCOR) consensus on science with treatment recommendations for pediatric and neonatal patients: pediatric basic and advanced life support. Pediatrics. 2006;117(5). Available at: www.pediatrics.org/cgi/content/full/117/5/e6955
29. Pocock SJ. When (not) to stop a clinical trial for benefit. JAMA. 2005;294:2228–2230
Evaluating Quality Improvement

Sheldon Greenfield, MD

Department of Medicine, Center for Health Policy Research, Irvine, California

The author has indicated he has no financial relationships relevant to this article to disclose.

The bloom is off the rose, at least partially. From the time they emerged in the 1940s as a substantial advance over available research designs, randomized, controlled trials (RCTs) quickly became the gold standard for testing medical interventions. However, it has long been acknowledged that for some interventions, particularly programs with multiple components or those intended for broad application or implementation when generalizability is of primary concern, RCTs for a variety of reasons either cannot be performed or are of less value than other research designs. A recent Institute of Medicine roundtable compendium encouraged movement toward increasingly sophisticated research designs other than RCTs to assess quality improvement or patient-safety initiatives.1–3

Rigorous evaluations of programs such as the California Asthma Among the School-aged (CAASA) project face the challenges and tensions inherent in testing the effectiveness of multiple-component interventions that are intended to be applied to highly disparate patient populations, widely varied clinical venues, and divergent geographic locations. In an extremely large, well-designed RCT that addressed effects of region, practice size, patient characteristics, and within-practice clustering, Homer et al4 found no effects of a quality improvement intervention program for children with asthma. The authors of that study specifically pointed to variability of program implementation and intensity and duration of the intervention as contributors to the null results. In pre-post, nonequivalent control-group studies of a national quality improvement project for adult patients with chronic disease, Landon et al5 found either modest improvements in the process of care, but no improvements in patients’ health outcomes,6 or no effects on quality of care at all.6 The authors also pointed to the variability or lack of intensity of the intervention after the initial sessions as contributing to the absence of an observed intervention effect.

In contrast, as shown in the October Pediatrics Electronic Pages, Fox et al7 performed a pre-post study of the impact of the CAASA project on the quality of care for children with asthma in 7 clinics distributed throughout the state of California. Rather than identifying a control group a priori, the authors defined a control group posthoc from those clinics that did not administer the quality improvement intervention according to study protocol. They further created the equivalent of a “dose-administration” variable, which was defined by the level of intensity of administration of their intervention or level of adherence to their study protocol.

Despite the acknowledged limitations of the authors’ study design, their findings that the CAASA project improved the quality of care for children with asthma are compelling for 3 reasons. First, all of the process and outcome measures of quality improved, often substantially, over baseline. On the basis of other empirical quality-of-care literature, magnitudes of the improvements shown in this study exceeded what would be expected from secular trends or regression to the mean. The uniformity and magnitude of these effects are particularly striking in light of the recent research showing mixed, if not disappointing, results from studies of quality improvement interventions.6–8 Second, although it

Abbreviations: RCT, randomized, controlled trial; CAASA, California Asthma Among the School-aged

Opinions expressed in these commentaries are those of the authors and not necessarily those of the American Academy of Pediatrics or its Committees.

www.pediatrics.org/cgi/doi/10.1542/peds.2007-2323
doi:10.1542/peds.2007-2323

Accepted for publication Sep 6, 2007

Address correspondence to Sheldon Greenfield, MD, Center for Health Policy Research, Department of Medicine, 111 Academy, Suite 220, Irvine, CA 92617. E-mail: sgreenfi@uci.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics.
does not meet standards for asserting causality, the evidence that Fox et al. presented documenting a strong association between level of adherence to their study protocol and improvement (adjusted for patient and site characteristics) supports the premise that observed improvements were related to the intervention. Their data are suggestive of a “dose response,” with greater improvements in quality related to more intensive administration of the intervention. Finally, and probably most important, is the plausibility of their findings. As a researcher and someone who has participated in quality improvement studies for many years, I am acutely aware of the difficulties encountered when trying to change patients’ and providers’ behavior. That it takes the intensity of a concerted program defined by the complete approach defined by the CAASA project is no surprise, particularly for poor and underserved patient populations, the health problems of which are known to be more difficult to address.

If we are to improve quality of care for the underserved, conducting large, expensive RCTs may not be the optimal strategy. Methodologic approaches such as that defined by Fox et al. should be considered for additional testing. If their results are replicable, costs of implementing and evaluating such interventions could be offset by savings in unnecessary emergency department visits and hospitalizations. Evaluations of cost-effectiveness of these more intensive approaches to quality improvement will be needed to determine their feasibility for broader implementation.

REFERENCES

COMMENTARY

The Changing Face of Preterm Birth

Mario Merialdi, MD, PhD, Jeffrey C. Murray, MD

Department of Reproductive Health and Research, World Health Organization, Geneva, Switzerland; Department of Pediatrics, University of Iowa Carver College of Medicine, Iowa City, Iowa

The authors have indicated they have no financial relationships relevant to this article to disclose.

In the October 2007 issue of Pediatrics, Edison et al. reported a remarkable association between low maternal serum cholesterol levels and preterm birth. Because preterm birth is the single largest factor worldwide in infant mortality and morbidity, this finding could have important consequences. The authors used an elegant approach with an existing cohort of women who had undergone a prenatal maternal serum screening analysis and provided a unique opportunity to correlate maternal serum cholesterol levels in early pregnancy with adverse birth outcomes. Because this finding was observed only in individuals of European ancestry, it will need to be replicated in other high-risk populations (eg, black Americans, preferably in a prospective manner). As the authors noted, the mechanism remains speculative but is biologically credible, because cholesterol is a precursor of placental progesterone, which is critical to the maintenance of pregnancy and is also a major component of plasma membranes in maternal decidual tissue. Although the association of low maternal serum cholesterol with preterm labor (as well as low birth weight) was unexpected, the finding opens the door to identification of a high-risk group of mothers in whom early interventions, including nutritional modifications or progesterone supplementation, might be effective in diminishing the subsequent occurrence of preterm labor.

Environmental components to preterm labor risk have long been recognized and include nutrition, infection, stress, trauma, and drug use. A genetic component is recognized but has been less well studied. A low maternal serum cholesterol level has multiple and overlapping etiologies, including maternal nutrition and genetic factors. It seems likely that the highest-risk group may well be those who have a combination of contributory dietary factors and a genetic predisposition to a low cholesterol level. Currently, there are active efforts by the World Health Organization, the National Institutes of Health, the Centers for Disease Control and Prevention, and the March of Dimes to better identify gene/environment etiologies for preterm birth and provide to the public information on the scope of the problem and the importance of recognizing early risk factors. Recent advances in genetics and environmental bioassays now make it practical to identify genetic and environmental interactions by using large sample collections and genetic tools such as genome-wide association analysis.

The application of these unique and powerful genetic techniques to common, complex traits such as preterm labor afford opportunities to prospectively identify high-risk populations and engage in preventative environmental interventions. The success in decreasing neural tube defects after the introduction of folic acid food fortification has provided a model approach. Given the well-recognized risks between cardiovascular disease and elevated cholesterol, interventions to raise cholesterol might best be conducted only in high-risk populations of women with low serum cholesterol levels early in pregnancy. This makes it imperative to identify the genetic underpinnings so that just as there is currently screening for Rh incompatibility, HIV status, and birth defects in early pregnancy, similar screening could be performed for mothers through simple assays of cholesterol levels and genetic predispositions to identify groups...
of women who would benefit from current or prepregnancy dietary counseling or, possibly, interventions with supplements of progesterone. Progesterone supplementation decreases the risk for preterm labor in some (previous preterm birth, short cervix) but not all high-risk groups so this might be a consideration in a low maternal serum cholesterol population. Preterm labor continues to be a major international health crisis, the frequency of which has seen a dramatic rise in developed countries over the last 2 decades, increasing the numbers of infants at risk for its many morbidity consequences. In addition, there is a vast difference in maternal and perinatal health status in high-versus low-resource settings. This disparity represents one of the starkest inequities of our time. Approximately 530,000 women and 3 million newborns die each year as a result of complications related to pregnancy and childbirth. Most of these deaths occur in developing countries among populations of lower socioeconomic status. However, even in wealthier countries disparities in pregnancy and newborn outcomes across socioeconomic strata persist. These alarming discrepancies indicate that perinatal health inequities are problems that transcend public health concerns alone and are a major social injustice.

Although research in maternal and perinatal health has made significant progress in recent years, most progress has been driven by the needs of health systems that operate in high-resource settings. This has resulted in the production of pregnancy- and childbirth-related interventions that translate poorly in low-resource settings and exacerbate the persistent gaps in women’s and infants’ reproductive health conditions around the world. The paucity of research funding and resultant efforts targeted at conditions that disproportionately affect women in low-resource settings has impeded the development of effective, affordable, and feasible preventive strategies. The research agenda for maternal and perinatal health in the future must be defined with a larger focus, one that targets the needs of all populations and especially those that are most vulnerable and in greatest need of affordable interventions. Successful implementation of this larger focus could lead to significant reductions in maternal and perinatal mortality, a goal that has yet to be achieved despite decades of international commitment. In addition, this new focus could have the added benefit of significantly reducing underlying causes of morbidity, disabilities, and associated health care costs in the more developed world. Results of the Edison et al study may well translate across socioeconomic strata and have wide application. These results could support the implementation of nutritional interventions among pregnant women to reduce the risk of preterm birth. Furthermore, it is possible that the effect of low cholesterol levels on preterm birth could be even larger among populations with inadequate nutritional status, which suggests that relatively simple, affordable, and culturally acceptable nutritional interventions could contribute substantially to reducing the risk of preterm birth among those populations most in need. The authors’ results also provide insights into the pathophysiology of parturition and suggest pathways to investigate for genetic contributors to preterm labor, as well. Although the results need confirmation, the Edison et al article provides critical data for beginning to understand and quickly apply information on the joint genetic and environmental mechanisms that cause this problem of major clinical and social importance.

ACKNOWLEDGMENTS

Dr Murray is supported by grants from the National Institute of Child Health and Human Development and the March of Dimes.

This work is dedicated to the memory of Dr Marcy Speer, who worked tirelessly to improve the lives of newborns.

REFERENCES

ABSTRACT
Although adolescent pregnancy rates in the United States have decreased significantly over the past decade, births to adolescents remain both an individual and public health issue. As advocates for the health and well-being of all young people, the American Academy of Pediatrics strongly supports the recommendation that adolescents postpone consensual sexual activity until they are fully ready for the emotional, physical, and financial consequences of sex. The academy recognizes, however, that some young people will choose not to postpone sexual activity, and as health care providers, the responsibility of pediatricians includes helping teens reduce risks and negative health consequences associated with adolescent sexual behaviors, including unintended pregnancies and sexually transmitted infections. This policy statement provides the pediatrician with updated information on contraception methods and guidelines for counseling adolescents.

INTRODUCTION
Pediatricians have an important role in adolescent reproductive health care. Their long-term relationships with patients and families allow them to help promote healthy decision-making around sexuality and include abstinence as a way to avoid the negative consequences associated with risky sexual behaviors. As advocates for the health and well-being of young people, pediatricians communicate their recommendation to adolescent patients to postpone sexual activity until they are ready, because any sexual activity for which the adolescent is ill prepared may have emotional, physical, and financial consequences. However, clinicians recognize that some of their adolescent patients are sexually active or will choose to become so. Recent studies indicate that, for some adolescents, even participating in formal programs that advocate abstinence and signing abstinence pledges do not result in abstinence behavior.1,2 Pediatricians can have an active role in encouraging their adolescent patients to use contraception to reduce the risk of unintended pregnancies and to prevent sexually transmitted infections (STIs). In previous publications, the American Academy of Pediatrics (AAP) has addressed issues of adolescent sexuality, unwanted pregnancy, STIs, and contraception.3 This policy statement provides the pediatrician with updated information on adolescent sexual behavior, which may lead to pregnancy, including guidelines for counseling adolescents about available methods of contraception. Current methods available are discussed, as are methods in development.

ADOLESCENT SEXUAL BEHAVIOR AND USE OF CONTRACEPTION
Reported contraceptive use by adolescents has increased in recent years. From 1991 to 2005, the percentage of sexually active high school students who reported using a condom the last time they had sexual intercourse increased from 46.2% to
Levels of reported sexual intercourse by adolescents in the United States decreased during the 1990s for both sexes after increasing for the previous 2 decades. The Centers for Disease Control and Prevention’s 2005 Youth Risk Behavior Surveillance Summary indicated that less than half (46.8%, down from 49.9% in 1999) of all high school students reported having had sexual intercourse in their lifetimes, and approximately one third (34.3%, down from 37.5% in 1991 and 36.3% in 1999) of all students reported having sexual intercourse during the 3 months preceding the survey and are considered currently sexually active.

Each year, almost 850,000 adolescent girls become pregnant. The adolescent pregnancy rate has dropped steadily over the past decade. As of 2004, it was estimated that approximately 41.2% of all pregnancies are to adolescents 15 to 19 years of age. Since 1991, the adolescent birth rate has declined by 33%, the lowest rate ever reported for the nation. The pregnancy rate for 15- to 17-year-olds has dropped by 43% to 22.1% of all pregnancies. Approximately 20% of abortions are in adolescents, although these rates continue to decrease.

Decreases in pregnancy rates are thought to reflect a decrease in reported rates of sexual intercourse and an increase in reported use of longer-acting, more effective contraceptive agents. Over the last decade, evaluations of curricula suggest that those with a comprehensive approach to sexuality education have been effective in improving sexual behaviors and, thus, may also contribute to this trend. Despite these declining rates of pregnancies and births, adolescent childbearing (22% of women report giving birth before age 20) is still more common in the United States than in other developed countries such as Great Britain (15%), Canada (11%), and France (6%).

Providing information to adolescents about contraception does not result in increased rates of sexual activity, earlier age of first intercourse, or a greater number of partners. In fact, if adolescents perceive obstacles to obtaining contraception and condoms, they are more likely to experience negative outcomes related to sexual activity. Two school-based studies that demonstrated a delay of onset of sexual intercourse used a comprehensive approach to sexuality education that included a discussion of contraception.

Race, ethnicity, age, marital status, education, income, requirements for confidential care, and fertility intentions have all been demonstrated to affect contraceptive choice. Trends in methods of contraception used by adolescents over the past 2 decades show an increase in oral contraceptive pill (OCP) use and an increase in male condom use. In recent years, the number of adolescents reporting OCP use has remained stable at approximately 18% to 20%. Use of injectable contraception by adolescents 15 to 19 years of age has increased from 0% to 13% between 1988 and 1995. A 9% decrease in contraceptive-failure–related pregnancies is attributed to the shift to longer-acting birth control methods.

Factors that contribute to lack of contraceptive use or inconsistent use include issues related to adolescent development, such as reluctance to acknowledge one’s sexual activity, belief that one is immune from the problems or consequences surrounding sexual intercourse or pregnancy, and denial of the possibility of pregnancy. Other important factors are lack of education and misconceptions regarding use or appropriateness of contraception. However, an adolescent’s level of knowledge about how to use contraception effectively does not necessarily correlate with consistent use. Adolescents may not use or may delay use of contraception for several reasons including lack of parental monitoring, fear that their parents will find out, ambivalence, and the perception that birth control is dangerous or causes unwanted adverse effects such as weight gain.

THE ROLE OF THE PEDIATRICIAN

Pediatricians should encourage abstinence and provide appropriate risk-reduction counseling regarding sexual behaviors. Ideally, counseling should include discussion about the prevention of STIs, education on contraceptive methods, and family planning services for the sexually active patient. Such discussion necessarily takes place within the context of an individual patient’s physical and emotional development as well as his or her social situation. Although pediatricians are optimally suited for such inquiry, we recognize that not every visit will allow the time required. The demands of comprehensive patient evaluation, counseling, and treatment are daunting, indeed, but are part of the ongoing education of teens and often other family members. This report is intended as a guide and, we hope, is helpful to busy clinicians.

When contraceptive services are provided in the pediatrician’s office, policies and procedures that address the provision of such services, including confidentiality, should be developed and then explained to families before the provision of such services is ever needed.

Counseling Adolescents About Contraception

Comprehensive health care of adolescents should include a confidential sexual history that should be obtained in a safe, nonthreatening environment through open, honest, and nonjudgmental communication with assurances of confidentiality. During the preadolescent years, the pediatrician can provide anticipatory guidance by discussing puberty and offering health education materials to both the youth and his or her family. At the onset of puberty, the patient’s history should include...
information on both the family’s and the patient’s attitudes and knowledge about sexual behaviors and the degree of involvement in sexual activity. General information may be offered or accessible to both the family and patient about methods of contraception and their uses. In addition, around this time, health maintenance visits should begin to include private, confidential time with the adolescent to establish rapport as well as assess degree of involvement in sexual activity. For sexually active adolescents who use contraception, the role of the health care professional is to educate and support compliance, to assist in managing adverse effects or, alternatively, to counsel the patient regarding a new contraceptive method as circumstances require, and to provide referrals and follow-up with periodic screening for STIs. Throughout adolescence, comprehensive sexuality education that includes discussion of abstinence, appropriate contraceptive use, and protection from STIs should be provided as part of healthy sexual development. When initiating any hormonal contraceptive method, the need for consistent protection against STIs (either male or female condoms) should be reinforced.

Confidentiality and Consent
The primary reason that adolescents may hesitate or delay obtaining family planning or contraceptive services is concern about lack of confidentiality. It is important for pediatricians to develop office policies that ensure patient confidentiality. State requirements and standards of practice should be reviewed, and the development of clear, concise, and standardized office protocols for confidentiality should be developed for staff, patients, and parents. These policies should include information and education regarding when confidentiality must be waived, guidelines for reimbursement of services, medical chart access, appointment scheduling, and information disclosure.

For those patients whose parents are unaware of their contraceptive use, it may be helpful to discuss with the adolescent patient how the contraceptive method will be consistently used in all circumstances. Consistent adolescent contraceptive use is often derailed during weekends away, family vacations, adolescents’ trips to stay with other relatives, and/or visits to noncustodial parents.

Sexual Responsibility
Pediatricians can help adolescents identify their own goals for safe and responsible sexual behavior, including reinforcing and supporting abstinence. The promotion of healthy and responsible sexual decision-making is one of the goals of counseling adolescents about contraception. Successful counseling requires a supportive and non-judgmental pediatrician who engages in effective dialogue, which includes skillful history taking, careful listening, and repetition of simple educational messages that contain essential information.

Sexual Decision-Making
Adolescents should be strongly encouraged to postpone or delay the initiation of sexual activity. For patients who are already engaged in sexual intercourse or who are contemplating having sexual intercourse, a discussion of contraceptive methods and prevention of STIs (including HIV and AIDS) is essential. Condom use should always be reinforced, and teens must be reminded that, for some STIs, condoms are not totally protective. Adolescents should be made aware, in a non-threatening and nonjudgmental manner, that although condom use is essential and may be life-saving, any individual who engages in sexual contact is at risk of contracting STIs that are transmitted through sexual contact, such as herpes simplex and human papillomaviruses, rather than body-fluid exchange, such as gonorrhea and trichomoniasis. Discussions should address and explore the adolescent’s reasons for becoming sexually active and the effect that sexual intercourse may have on relationships with peers, parents, and significant others. Clinicians may also find it useful to explore with the adolescent how he or she believes the sexual experiences will change his or her own self-image. Adolescent sexual decision-making has emerged in recent studies as a complex interplay between an adolescent’s perception of peer-group expectations, personal self-image, values, and desires and media influences. However, a caring, nonjudgmental yet informative, nonparental adult can wield substantial influence in teens’ sexual decision-making; teens cite lack of such a person as a missing key feature of sexuality education. Pediatricians, therefore, may have some influence in adolescent sexual decision-making and are especially well-positioned to assess risk-taking behaviors in the area of sexuality.

Adolescents with Disabilities
The issue of contraception for adolescents with chronic illness or disability is often forgotten. An estimated 10% to 20% of children and adolescents experience a disability or chronic illness by the age of 20 years. Recent data from the National Longitudinal Study of Adolescent Health has shown that physically disabled adolescents are as sexually experienced as adolescents without disabilities. Attitudes about contraceptives as well as sexuality education and counseling needs within this population should not be overlooked. A list of additional resources for clinicians who desire more information about contraception for adolescents with chronic illness and/or disability is included at the end of this statement.
METHODS OF CONTRACEPTION

Numerous reviews and protocols for prescribing and managing contraception are available. The following section focuses on the appropriateness of various contraceptive methods available for adolescents. The pediatrician should emphasize the need for STI prevention as well as contraception with each patient at each visit.17,38

Abstinence
Abstinence is the most effective means of birth control and prevention of STIs and is a viable strategy in the clinician’s toolkit for reducing unintended pregnancy and achieve reduction in STI rates. Abstinence education generally focuses on delaying the initiation of adolescent sexual activity until marriage or adulthood. Many schools have adopted abstinence-dominant or abstinence-only education programs for school sexuality curricula. To date, the evidence regarding the efficacy of such interventions in the reduction of risky sexual behaviors, including risk for STIs, has not been proven.14,39

No data have directly examined how well abstinence counseling works to reduce an individual’s pregnancy and STI risk. In practice, many adolescents who intend to be abstinent often fail and have sex. A longitudinal analysis of teens and virginity pledges compared pledgers to nonpledgers and found that at 6-year follow-up that 88% of pledgers reported experiencing premarital sex and had STI rates that, statistically, were no different from those of nonpledgers.2 A recent article provides some practical tips for abstinence counseling within an office-based setting using a comprehensive perspective including motivational interviewing.40

Several published studies and evaluations have suggested that comprehensive sexuality education is an effective strategy for helping young people delay initiation of sexual intercourse. In addition, research has shown that these programs do not hasten the onset or frequency of sexual intercourse and do not increase the number of partners that sexually active teens have.15

There is some consensus that sexuality education and interventions with some abstinence-based or “abstinence-plus” curriculum components are most effective when targeted at younger adolescents before they become sexually active.1,41 Some recent studies demonstrated the importance of youth, parent, physician, and education partnerships for the prevention of health risk behaviors such as early initiation of sexual intercourse.13,42 One study illustrated that an abstinence-only curriculum had no significant impact on the initiation of sex, the frequency of sex among those students who had ever had sex, or the number of sexual partners among those who had ever had sex. Two other studies produced similar results.1 The AAP supports a comprehensive approach to sexuality education for adolescents. Abstinence should play a part in any comprehensive discussion of sexuality, and resources should be made available for adolescents who feel pressured, but prefer not, to engage in sexual activity.

For some adolescents, abstinence may be a difficult choice. Adolescents who choose to abstain from sexual intercourse should be encouraged and supported by their parents, peers, and society (including the media) and especially by their pediatricians. Adolescents need to know about other contraceptive options before (or if) they decide to have intercourse.

Male and Female Condoms
The male condom is a mechanical barrier method of contraception. The failure rate at the end of first-year use for the male latex condom is 3% with perfect use and as much as 14% with typical use.43 Latex condoms significantly reduce the transmission of some STIs and, therefore, should be used by all sexually active adolescents regardless of whether an additional method of contraception is used. Male condoms have several other advantages for adolescents, including involving males in the responsibility of contraception, easy accessibility and availability to minors, use without a prescription, and low cost.44 Polyurethane condoms can be used by adolescents with a documented latex allergy; however, latex condoms are preferred, because they have a higher efficacy rate with typical use than polyurethane condoms.43 Some adolescents may have local reactions to condoms that have been pretreated with spermicide and should be counseled that condoms without these agents are also available.

Nonoxynol-9 is the only chemical agent in spermicidal products available in the United States; there are nonspermicidal hypoallergenic lubricants available over the counter. Only water-based lubricants may be used with latex condoms, and both water- and oil-based products may be used with polyurethane condoms.44 Currently, there is a general movement away from products with nonoxynol-9 because of concerns that use increases risk of genital ulceration and irritation, which may facilitate the acquisition of STIs.44 Condom use reported at most recent intercourse by females was 54% and by males was 71%, which is an increase in the last decade.45 Surveys of high school students over the last decade indicate that condom use has increased, with condom use at last intercourse increasing from 46.2% in 1991 to 62.8% in 2005.4

The female condom, another barrier method of contraception, provides contraceptive efficacy in the same range as other barrier methods, such as the diaphragm and cervical cap (with typical use).45 One trial of the most widely available female condom on the market yielded a failure rate of 0.8% with perfect use and between 12% and 15% with typical use.46 The female condom also helps protect against STIs. Adolescents’ concerns about using a female condom include difficulty of insertion, higher cost than male condoms, and appearance and noisiness of the device. Female adolescents
have reported that the female condom could be useful if their male partners did not want to use a condom. Further education on using the female condom is needed for both genders. For adolescents who already use male condoms, it is important to market the female condom as an alternative contraceptive choice, because male and female condoms should not be used simultaneously.49 Male condoms are preferred over female condoms because of higher efficacy rates of preventing pregnancy and STIs and lower cost.

Vaginal Spermicides

Vaginal spermicides are a chemical barrier method of contraception applied intravaginally through a variety of forms: gel, foam, suppository, or film. Spermicides consist of 2 components: a formulation (the gel, foam, suppository, or film) and the chemical ingredient that kills the sperm (eg, nonoxynol-9). As with any barrier method, the effectiveness of spermicides depends on consistent and correct use. The combination of vaginal spermicide and condoms is a very effective means of contraception for adolescents, because it provides effective prevention of pregnancy, reduces the risk of contracting an STI, is available without a prescription, and is inexpensive.48

There has been a question as to whether use of nonoxynol-9 alone provides adequate protection against STIs and HIV. In high doses, nonoxynol-9 can irritate the vaginal lining, which makes young women more susceptible to HIV transmission. The Centers for Disease Control and Prevention have concluded that women should be discouraged from using nonoxynol-9 alone for STI and HIV protection, because 1 study found that a product containing nonoxynol-9 did not protect against HIV infection and may have caused an even greater likelihood of transmission as compared with a vaginal lubricant.49,50 Use of spermicide alone is not advocated as a contraceptive method; condoms must be used in conjunction with vaginal spermicides for protection against STIs.

Oral Contraceptives

OCPs are a reliable, effective method for the prevention of pregnancy, are available only by prescription in the United States, and are the most popular method of prescribed contraceptive among adolescents.21 Of the 2.7 million adolescent women who use contraceptives, 44% rely on the pill.51 The Youth Risk Behavior Surveillance Summary reported that in both males and females who had sexual intercourse during the 3 months before the survey, the percentage who used birth control pills to prevent pregnancy during last sexual intercourse was 17.6% in 2005, down from 20.8% in 1991.4

Three forms of OCPs are currently available: the fixed-dose, monophasic combination (each tablet contains the same dose of estrogen and progestin; the pharmacologic dose (the triphasic and biphasic packs that contain varying doses of estrogen and progestin); and the mini-pill (which contains progestin only). Many of the newer forms of birth control pills have a low dose of estrogen (20–35 μg) and contain new forms of progestin. These low-dose pills are typically the “first-line” therapy for OCP initiation. There is theoretic potential for lowered efficacy of low-dose OCPs in patients who are taking some medications. Some common medications that increase the metabolism of synthetic steroids by increasing conjugation in the gut and enzyme induction in the liver are listed in Table 1.52–54 In this clinical situation, prescription of OCPs that contain 50 μg of ethinyl estradiol or switching to a hormonal method that avoids first-pass metabolism, such as injectable progestin, may be indicated; efficacy of transdermal or intravaginal contraceptives with these medications is not known. Generally, the standard 28-day pack of pills (21 days of hormone and 7 days of placebo) is prescribed for teens, and daily compliance is encouraged, particularly over the 21 days of hormone-containing pills to maximize efficacy and minimize bleeding irregularities.55 The 21-day packs, if available, are better for adolescents who are taking OCPs in continuous or extended cycles.

The US Food and Drug Administration (FDA) recently approved a monophasic 30-μg ethinyl estradiol/0.15-mg levonorgestrel pill for extended cycling called Seasonale (Barr Pharmaceuticals, Woodcliff Lake, NJ). This formulation provides 84 days of continuous hormonally active pills followed by 7 days of placebo. This formulation may be particularly appropriate for adolescents with medical conditions such as anemia, severe dysmenorrhea, endometriosis, dysfunctional uterine bleeding, or Von Willebrand and other bleeding diatheses and adolescents who prefer amenorrhea.56 In addition, adolescents who frequently miss OCPs may have lower failure rates when using continuous or extended regimens of OCPs with shorter or no placebo intervals.

<table>
<thead>
<tr>
<th>TABLE 1 Medications That Decrease OCP Efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibiotics</strong></td>
</tr>
<tr>
<td>- Rifampin (Rifadin, Rimactane)</td>
</tr>
<tr>
<td>- Anticonvulsants</td>
</tr>
<tr>
<td>- Felbamate (Felbatol)</td>
</tr>
<tr>
<td>- Ethosuximide (Zarontin)</td>
</tr>
<tr>
<td>- Primidone (Myidone, Mysoline)</td>
</tr>
<tr>
<td>- Phenytoin (Dilantin, Phenytek)</td>
</tr>
<tr>
<td>- Carbamazepine (Tegretol)</td>
</tr>
<tr>
<td>- Oxcarbazepine (Trileptal)</td>
</tr>
<tr>
<td>- Topiramate (Topamax)</td>
</tr>
<tr>
<td><strong>Antidepressants</strong></td>
</tr>
<tr>
<td>- St John’s wort</td>
</tr>
<tr>
<td>- Antifungal agents</td>
</tr>
<tr>
<td>- Griseofulvin (Fulvicin, Grifulvin, Gris-PEG)</td>
</tr>
<tr>
<td><strong>Anti-HIV drugs</strong></td>
</tr>
<tr>
<td>- Protease inhibitors</td>
</tr>
<tr>
<td><strong>Nonnucleoside reverse transcriptase inhibitors</strong></td>
</tr>
</tbody>
</table>

PEDIATRICS Volume 120, Number 5, November 2007 1139
The noncontraceptive benefits of OCP use include improvement in acne and decreased menstrual cramping, pain, blood loss, and ovarian cysts. OCP use that exceeds 3 years provides significant protection against endometrial and ovarian cancers. Overall, observational data indicate that OCP use does not increase risk of breast cancer. Adverse effects include nausea, breast tenderness, headaches, and breakthrough bleeding. OCPs are one of the best-studied medications ever prescribed and are a safe option throughout a woman’s reproductive years, because the method is completely reversible and has no negative effect on long-term fertility.

OCPs have a failure rate of 0.1% when used perfectly. However, failure rates range between 5% and 8% with typical use and for adolescents may reach 15% to 26% because of noncompliance. Adolescents may have difficulty complying with OCPs because of forgetfulness, attempts to hide contraception from parents, and inconsistency of sexual relations, among other reasons. The National Survey on Family Growth reported that as many as 42% of adolescents 15 to 19 years of age missed 2 or more pills in a 3-month period. Adolescent compliance with OCP use may be enhanced by appropriate patient education and problem-solving techniques, which includes careful instruction regarding the use of OCPs; anticipatory guidance about adverse effects and their management; a discussion of correct pill usage (including when the first pill should be taken during the menstrual cycle or what to do if a pill or pills are late or missed); use of emergency contraception; and frequent follow-up and monitoring. Patients should also be encouraged to use condoms in conjunction with OCPs to provide protection against STIs and additional pregnancy prevention. In addition, when possible, involving the patient’s mother can greatly enhance compliance with pill taking.

OCPs have few contraindications in healthy female adolescents. Estrogen-containing OCPs are contraindicated for those with a history of thromboembolism or thrombophilia (ie, factor V Leiden mutation or protein C, protein S, or antithrombin III deficiencies); cyanotic heart disease or pulmonary artery hypertension; systemic lupus erythematosus associated with antiphospholipid antibody syndrome or renal disease, particularly associated with hypertension; or hepatic dysfunction. Patients who are taking anticonvulsant medications and HIV medications need to be counseled carefully and may be encouraged to use injectable progestins (Table 1).

Adolescents need not receive a complete gynecologic examination by the pediatrician before initiating OCPs or any other hormonal contraceptive method. In most circumstances, the pelvic examination may be deferred and OCPs may be prescribed if the patient is healthy, is not pregnant, and has no contraindications to taking the pills. An inspection of the external genitalia and either a urine screen or vaginal swab for STIs may be substituted for a pelvic examination as a screening for initiation of contraceptive use. A pelvic examination is indicated for most situations in which abdominal pain is part of the presenting complaint in a sexually experienced adolescent. Sexually active female adolescents should be screened for STIs, especially chlamydia, at least annually and preferably with each new sexual partner. Guide- lines from the American Cancer Society and the American College of Obstetricians and Gynecologists recommend initiation of Papanicolaou (Pap) test screening within 3 years of first intercourse (whether consensual or nonconsensual) or by 21 years of age.

Injectable Hormonal Contraception

Depot medroxyprogesterone acetate (DMPA) injection is a long-acting progestin that is given every 12 weeks (11–13 weeks) as a single 150-mg intramuscular dose. This method of contraception, also known by the brand name Depo-Provera (Pfizer, New York, NY), is highly effective in preventing pregnancy. In the first year of use, the probability of becoming pregnant is approximately 0.3%. Available since 1992 in the United States, some experts believe that the use of this method since 1992 among adolescents who are at high risk of becoming pregnant is one factor responsible for the declining rates of adolescent pregnancy in the United States. This method is convenient for women who do not want to have to remember to take their pill each day, cannot use the patch, or cannot use a contraceptive at the actual time of intercourse. Other advantages include lack of estrogen-related adverse effects and, similar to OCPs, protection against endometrial cancer and iron-deficiency anemia.

The major disadvantage of this contraceptive method for adolescents is menstrual cycle irregularities (present for nearly all patients initially), the need for intramuscular administration every 11 to 13 weeks, and potential adverse effects including acne, weight gain, headaches, and bloating. A new formulation, which is administered subcutaneously, contains 104 mg of medroxyprogesterone acetate (Depo-Subq Provera 104 [Pfizer]), and is given on the same dosing schedule as the intramuscular formulation, is now available. The subcutaneous route makes home administration of Depot-Provera possible, although there have been no studies of home use in the adolescent population. The lower dose could decrease suppression of pituitary function and ovarian estradiol production, although no conclusive data are yet available to indicate such an effect. Two large open-label phase-3 studies have found subcutaneous DMPA to be equally effective as intramuscular DMPA; however, the irregular uterine bleeding that many patients complain of after initiating the drug also accompanies subcutaneous use. As with the intramuscular route, this adverse effect largely resolves over the
first year of use: amenorrhea increased from 26% of patients in month 3 of use to 55% during month 12.66

In addition to uterine bleeding irregularities, DMPA use over a prolonged period is associated with a delayed return to fertility, typically 9 to 18 months, while the endometrial lining returns to its pre-DMPA state and ovulatory function returns. Both subcutaneous and intramuscular DMPA show similar delays to fertility after injection.67 However, for adolescent patients, such a delay does not usually pose a major deterrent to using this method. Both intramuscular and subcutaneous DMPA may be safely recommended for adolescents who have chronic illnesses (eg, seizures, sickle cell disease), are lactating, or are at risk of estrogen-related complications.64

Pediatricians should discuss potential adverse effects. Studies have shown that patients are more likely to continue DMPA use if they are counseled about potential irregular bleeding before their first injection, but these studies did not target adolescents specifically.68 Clinicians must also ensure that the patient is not pregnant at the time of the initial injection and at each injection that occurs at an interval greater than 12 weeks.

Because DMPA suppresses circulating estradiol concentrations, it causes reductions in bone mineral density (BMD), which has generated some concern regarding the long-term effects. A prospective cohort study of adolescents aged 12 to 18 years found that BMD decreased 3.1% after 2 years of DMPA use, whereas BMD increased 9.5% in the controls who were using no hormonal method of contraception.69 Some other studies have indicated an adverse impact on biochemical markers of bone formation and resorption as well as the decreased BMD.70–72 In response to these concerns, the FDA issued a “black-box” warning regarding the risk of decreased BMD among DMPA users in November 2004.73 Currently, the warning recommends limiting the use of DMPA to 2 years and using DMPA as long-term contraception only if other methods are inadequate. The warning also emphasized the lack of certainty regarding peak BMD attained later in life among users of DMPA, but experts think such a restriction may be unwarranted, especially for patients with no other alternatives for contraception. A recently published study of teens and young adult women documented complete recovery of BMD after DMPA use, thus offering some degree of reassuring about use not affecting long-term skeletal health of adolescent patients.74 In addition, an increased incidence of fractures has not been reported in adolescents using DMPA.

It is important to consider other risk factors for osteoporosis and to tailor counseling and recommendations to each patient. Factors such as small body habitus, chronic alcohol or tobacco use, eating disorders, or illness that necessitates chronic use of corticosteroids may lead a clinician to recommend against DMPA use more strongly. All patients should be encouraged to include foods and/or supplements to ensure intake of at least 1300 mg of calcium each day along with 400 IU of vitamin D, to participate in weight-bearing exercise regularly, and to stop smoking as important measures to promote skeletal health. Using supplemental estrogen has been observed to prevent loss of BMD in 1 study of teens, whereas the use by teens of antiresorptive medications prescribed for postmenopausal women is definitely not recommended.75 As with all hormonal methods of contraception, condoms should be used in conjunction with DMPA for protection from STIs.

Another injectable hormonal contraceptive (known by the product name Lunelle [previously manufactured by Pharmacia & Upjohn, Kalamazoo, MI]) combined estrogen and medroxyprogesterone acetate. Lunelle was made available in the United States after confirmation of safety in clinical trials in the United States and internationally. However, Lunelle was voluntarily withdrawn from the market by its manufacturer in September 2002 because some doses may not have contained enough hormone to prevent pregnancy. Women who used Lunelle required monthly clinic visits for the injection of medications. Adverse effects were similar to those of Depo-Provera and included weight gain, menstrual irregularity, headaches, and breast tenderness, although adverse effects were fewer than in trials with DMPA alone.76–77 A general acceptance of and overall satisfaction with Lunelle by women in clinical trials suggested that this method was widely accepted, but its return into the market is not expected in the future.

Progestin Implants

Levonorgestrel implants, also known by the brand name Norplant (previously manufactured by Wyeth-Ayerst Laboratories, St Davids, PA), were highly effective long-acting progestin-only contraceptives that provided pregnancy prevention for up to 5 years. These implants required insertion of subcutaneous polymeric silicone capsules into the upper arm by a trained health care professional. The 6-rod Norplant system was the first progestin implant available in the United States but has been permanently removed from the US market.78 Implanon (Organon USA, Roseland, NJ), a single-rod implant that contains etonogestrel, the active metabolite of desogestrel, has been used in Europe since 1998 and is now available in some areas of the United States. Highly effective (in clinical trials, no unintended pregnancies were reported in ~73 000 cycles), Implanon may remain in place for 3 years, but it is associated with irregular bleeding in many users, especially during the first year of use.63

Levonorgestrel implants are ideal for adolescents who desire an extended length of protection, feel unable to remember to take OCPs, or have already had 1 preg-
nancy. It is also an excellent contraceptive option for females who may have difficulty remembering to use a contraceptive on a daily basis or at the time of intercourse. The major disadvantages for use in the adolescent population include high initial cost and potential adverse effects such as breakthrough bleeding and headaches. The drugs listed in Table 1 also impair the efficacy of levonorgestrel implants. The difficulty of removal of the implant, in combination with these other disadvantages, made Norplant an unpopular form of contraception for adolescents, and although Implanon is easier to remove, it shares many of Norplant’s adverse effects. In addition, condoms must be used in conjunction with progestin implants for protection against STIs.63

Other Combined Hormonal Contraceptive Methods (NuvaRing and Ortho Evra)
The vaginal ring (NuvaRing [Organon USA]; 15 μg ethinyl estradiol/120 μg etonogestrel) is a round, flexible device that measures 54 mm in outer diameter and 4 mm cross-sectionally; it is inserted in the vagina and stays in place for 3 weeks, with removal for 1 week to induce menstruation followed by insertion of a new ring. This soft silicone vaginal ring releases both estrogen and progestin hormones that protect against pregnancy for 1 month. The ring has been shown to have greater than 99% efficacy when used correctly by adult women. However, trials with adolescent populations have not been conducted. Compliance with the ring is high, and few adverse effects are experienced. Adverse effects would be the same as other combined hormonal methods, which include breast tenderness, headaches, nausea, and some breakthrough bleeding/spotting and an increased risk of the more serious condition of thrombotic events; local adverse effects may include vaginal symptoms of discharge, discomfort, and device problems.55

The combination hormonal transdermal adhesive skin patch (Ortho Evra [Ortho-McNeil Pharmaceutical, Raritan, NJ]) can be applied to the abdomen, upper torso, upper outer arm, or buttocks weekly by using 1 patch for each of 3 weeks in a row, followed by 1 week off the patch, during which a withdrawal bleed usually occurs. While in place, the 4.5-cm³ contraceptive patch delivers 150 μg of norelgestromin and 20 μg of ethinyl estradiol daily. Efficacy rates from 1 study suggested that the overall annual probability of pregnancy was 0.8%, whereas the method failure probability was 0.6%, similar across age and racial groups.55

One study reported that women who use the patch are no more likely to become pregnant than women who use a combination OCP. At least 1 study indicated a higher rate of local adverse effects with adolescents who use the patch than with older patients; these effects include patches dislodging as well as irritation and hyperpigmentation.79,81 Very concrete counseling regarding patch placement with adolescent patients, and perhaps even demonstration of initial placement, is helpful.81

Although compliance with using the patch is improved compared with OCPs, the risk of pregnancy with correct use of the patch was higher for women who weigh more than 198 pounds (0.9% in first 12 months of use) compared with women who weigh less (0.3%).82,83 Obviously, the risks of pregnancy must be discussed as methods are considered; a pregnancy rate of 1% at the end of 1 year in a patient who weighs more than 200 pounds, refuses other methods, and chooses to remain sexually active is a more acceptable alternative than a pregnancy risk of 85% with no protection. For issues related to compliance, the added value of the patch should be considered. It has demonstrated increased compliance, which results in fewer contraceptive failures. Other possible adverse effects of combined hormone methods include temporary irregular bleeding, temporary breast discomfort, weight gain or loss, and nausea.

The most concerning possible adverse effect of transdermal contraception or any combined hormone method is risk of thrombotic events. In the large clinical trials of the transdermal contraceptive patch, 1 case of nonfatal pulmonary embolism occurred during use of the patch, and 1 case of postoperative nonfatal pulmonary embolism was reported. Recently, the higher bioavailability of estrogens delivered transdermally has prompted Ortho McNeil Pharmaceutical to issue a warning specifically related to risk of thrombotic events in Ortho Evra users.84 An increased risk of venous thromboembolic disorders has been associated with the use of combination hormonal oral contraceptives compared with nonusers; generally, increases in rates by a factor of 3 to 6 have been reported in studies that evaluated healthy young women who had no other risk factors. A review of studies compared the risk of nonfatal venous thromboembolism (VTE) among different OCPs. The authors found that users of OCPs containing desogestrel, a third-generation progestin, had an increased risk compared with users of OCPs containing levonorgestrel.85 Thus, a baseline risk for VTE of 1 per 10 000 person-years is increased to 3 to 4 per 10 000 person-years during the time when oral contraceptives are being used.85 The risk is much greater in those over age 35 and those who smoke, especially if cigarette use equals or exceeds 15 cigarettes daily. Although smoking should be discouraged for teens and young adults, smoking and use of combined hormone methods of contraception are not contraindicated in this age group.86 Whether the patch places teens and women at increased risk of VTEs compared with combined hormone OCPs is not known, because 2 different studies with different methodologies had different outcomes.87 No studies to date have directly examined whether the patch increases the risk of
Intrauterine Devices
Intrauterine devices (IUDs) are inserted into the uterus and release hormones, ions, or enzymes that prevent sperm from fertilizing the ova or prevent implantation. The effectiveness of IUDs is influenced by several factors, including size of the IUD surface area and the type of IUD used. When used appropriately, IUDs are generally safe, effective methods of contraception with a failure rate of less than 1%. Condoms must be used in conjunction with IUDs for protection against STIs. IUDs have previously not been recommended for adolescents; risk of infection in teens (who often have multiple partners or are serially monogamous) and liability concerns (a patient who has not conceived before using an IUD may attribute future infertility to IUD use) have contributed to clinicians’ reluctance to prescribe this method for adolescent patients. IUDs have not been shown to affect fertility in the absence of infection; however, STI rates in adolescent populations are certainly cautionary. In some cases, however, an IUD may be appropriate for an adolescent who already has children and is taking precautions to protect against STIs. Mirena (Bayer, Montville, NJ), a newly developed IUD that contains the progestin levonorgestrel, gradually releases the progestin over an effective period of 5 years and has a failure rate of 0.3%. This IUD may be particularly useful for adolescents with severe menorrhagia and dysmenorrhea, as has been shown in adult women. Also available is the copper IUD called ParaGard, which releases a small amount of copper that kills or immobilizes sperm before they can fertilize an egg. The ParaGard can be removed at any time but should be replaced after 10 years.

Diaphragm and Cervical Cap
The diaphragm and cervical cap are barrier methods of contraception that also require use of a spermicide. Diaphragms are flexible latex cups that are inserted into the vagina before intercourse and must remain in place for 6 hours after intercourse. Cervical caps are latex or silicone cups with a firm rim that adhere to the cervix and provide continuous contraceptive protection for up to 48 hours. Because of risk of toxic shock syndrome, caps should not remain for more than 48 hours. Consistent, correct use of these methods is critical for achieving a high rate of effectiveness. The failure rate of the diaphragm with perfect use is 6% and with typical use is 20%; for the cervical cap, the failure rate is 26% with perfect use and 40% with typical use. These contraceptive methods may not be feasible for some adolescents, because they require a prescription and visit with a health care professional for a fitting, because 1 size does not fit all. The adolescent must also be comfortable and skilled with insertion. Incidence of urinary tract infection increases over baseline with both diaphragms and cervical caps; in addition, condoms must be used in conjunction with these devices for protection against STIs.

Withdrawal
The withdrawal method, which involves the male partner’s attempt to withdraw the penis before ejaculation, is still widely used by adolescents in sexual relationships. Adolescents should receive counseling that emphasizes the high failure rate of withdrawal for pregnancy prevention. On average, of every 100 women whose partners use withdrawal, 19 will become pregnant during the first year of typical use. It is important to stress that preejaculatory fluid can contain enough sperm to cause pregnancy. Pregnancy is also possible if semen or preejaculate leaks out onto the vulva. In addition, providers should stress that this contraceptive method does not provide protection against STIs.

Fertility Awareness and Other Periodic Abstinence Methods
Using fertility-awareness methods as a contraceptive option depends on several factors and requires a strong knowledge of the menstrual cycle and reproductive fertility. This method involves the identification of fertile days within each menstrual cycle when intercourse is most likely to result in pregnancy. Couples can abstain during the fertile times of a woman’s cycle or use a combination of either barrier or withdrawal methods. As many as 25% of users of these methods will experience an unintended pregnancy within the first year of use, with some estimates of the pregnancy rate even higher. To optimize method efficacy, users of this method should track their menses on a calendar for 3 months while also checking and recording their basal body temperature daily and should check their cervical mucus consistency to track when they ovulate. Pediatricians should be prepared to teach adolescents about the menstrual cycle but should emphasize that ovulation may not be predictable in the first few year(s) after menarche. Thus, abstinence or more reliable methods should be recommended for adolescents. In addition, health care professionals should stress that this contraceptive method provides no protection against STIs if no barrier methods are used during periods of sexual activity.
Emergency Contraception

Emergency contraception can be administered in 2 ways: by orally administering hormones or by inserting a copper-releasing IUD. An IUD can be inserted to prevent pregnancy up to 5 days after unprotected intercourse but is usually not recommended for adolescents (see IUD section).

The most commonly prescribed and best-studied methods of emergency contraception are the combined estrogen-progesterin (also called the Yuzpe regimen) and progestin-only regimens. There is now only 1 dedicated product for emergency contraception: Plan B (DuraMed Pharmaceuticals, Pomona, NY). Plan B, a progestin-only regimen that contains levonorgestrel, is widely available as 2 hormone pills that are taken within 72 hours of unprotected intercourse. The most recent data support extending the time limit of use to 120 hours after unprotected intercourse; however, emergency contraception’s efficacy diminishes as hormonal administration becomes more remote from the unprotected intercourse event.92–94 Adolescent patients especially should be counseled that Plan B is 90% effective if used within 24 hours, 75% effective if used within 72 hours, and approximately 60% effective if used within 120 hours. The Plan B regimen can now be simplified to give both tablets at one time without sacrificing efficacy or resulting in more adverse effects.94 Combination OCPs may be used for emergency contraception when Plan B is not readily available; the dose depends on the specific product chosen.95 A recent study found that combination OCPs with progestin norethindrone can also be used effectively for emergency contraception. This study found that even a single dose of the oral contraceptives or combined hormone method was effective for emergency contraception.94 Adverse effects may include nausea, vomiting, and changes in the menstrual cycle during the month of use. The progestin-only regimen is generally preferred, because it is more effective and causes fewer adverse effects.95,98 Overall, emergency contraceptives reduce the risk of pregnancy after unprotected sex by at least 74%.99,100

Most women who need emergency contraception can use it safely. If the patient or practitioner suspects pregnancy, a pregnancy test can be administered; however, pregnancy testing before emergency contraceptive use is not necessary. It is important to note that emergency contraception does not cause abortion and it is not teratogenic if taken in early pregnancy. Women who are already pregnant should not use emergency contraceptives because they are ineffective at terminating established pregnancies; however, using them inadvertently will not have an adverse effect on the fetus.99 Six studies have found that providing emergency contraception in advance increases the likelihood of women using it when it is needed and does not increase sexual or contraceptive risk-taking behavior.101–106

As the AAP states in its policy statement on emergency contraception, reduction of unintended pregnancy is best achieved by strategies that include developing and implementing programs to help delay and reduce sexual activity and increasing the use of effective contraceptives.93 However, the AAP continues to support improved availability of emergency contraception to adolescents and advocates clinicians’ consideration of advance emergency contraception prescription to sexually active adolescents, recognizing that in some cases, emergency contraception may be quite valuable in preventing unintended pregnancy and that emergency contraception is most effective when used soon after unprotected intercourse.95 Recently, the FDA approved over-the-counter access for Plan B for women 18 years and older, but Plan B still requires a prescription for those younger than 18 years.107 In view of the potential value of emergency contraception, pediatricians should inform adolescents about the availability of emergency contraception; however, it should not be advocated as a routine method of contraception.

Newer Forms/Formulations of Contraception

The FDA recently approved the first chewable OCP, Ovcon 35 (Bristol Myers Squibb Company, Princeton, NJ), a spearmint-flavored, 28-day regimen pill that contains the same hormones used in standard OCPs. The FDA recently approved the first chewable OCP, Ovcon 35 (Bristol Myers Squibb Company, Princeton, NJ), a spearmint-flavored, 28-day regimen pill that contains the same hormones used in standard OCPs. Women who chew the pills instead of swallowing them should drink 8 oz of liquid afterward to ensure that the full dose reaches the stomach.108 Another method recently approved by the FDA is the FemCap, a soft silicone dome that covers the cervix. FemCap will be available by prescription in 3 sizes and is designed to last 48 hours per use.89,109 New forms of contraception for males are also being studied, including an implantation system similar to Norplant, weekly and monthly hormone injections, and a contraceptive patch.110 A progestin-only vaginal ring is being developed, and Norplant II (a 2-rod system as opposed to the 6-rod system in Norplant) is awaiting FDA approval. Condoms must be used in conjunction with these new forms of contraception for protection against STIs.

COMPLIANCE AND FOLLOW-UP

Frequent follow-up is important to maximize compliance for all methods of contraception, to promote and reinforce healthy decision-making, and to screen periodically for risk-taking behaviors and STIs. Follow-up visits should include periodic examinations, reassessment for contraception method, STI surveillance, and cervical cytologic screening (Papanicolaou test) when appropriate. The timing and frequency of reassessment will vary depending on the contraceptive method. In general, sexually active adolescents should have annual STI screening with consideration for repeat screening for chlamydia 3 to 6 months after a positive test result and
treatment and/or with each new partner. Regularly scheduled visits need to occur to assess contraceptive issues such as use, compliance, adverse effects, and complications. Adolescents should receive ongoing support, personal guidance, and reinforcement to enhance effective and consistent contraceptive use, parental support (when possible), and couples counseling or the opportunity for couples interaction with the health care professional. In addition, condom use at each sexual intercourse must be advised and reinforced at every visit.

RECOMMENDATIONS

1. Pediatricians should encourage sexual abstinence as part of comprehensive sexuality education and services offered to their adolescent patients.

2. Pediatricians should be prepared to offer confidential, nonjudgmental education and risk-reduction counseling around issues of sexuality for adolescent patients, including teens with chronic illnesses and/or disabilities.

3. Pediatricians should be aware that extensive information regarding contraceptive choices and decisions for adolescents with chronic illness or disability are available in references and texts on adolescent medicine (see “Additional Resources”).

4. Pediatricians should update each patient's sexual history regularly to counsel about and determine risk of STIs as well as needs for contraceptive initiation and management.

5. Time to counsel, educate, and solve problems regarding contraceptive needs and/or management needs to be a part of any given visit, or arrangements need to be made for a separate visit for contraceptive follow-up.

6. Pediatricians should encourage the consistent and correct use of latex condoms with every event of sexual intercourse.

7. Pediatricians should know that it is appropriate to prescribe contraceptives without a “first pelvic examination,” but screenings for STIs, especially chlamydia infections, should not be delayed.

8. Pediatricians should ensure access to basic contraceptive services for their teen patients either within their office setting or by referral to appropriate services and/or sites.

9. Pediatricians who offer contraceptive services to adolescents should provide appropriate follow-up to ensure compliance and monitor for adverse effects and complications.

COMMITTEE ON ADOLESCENCE, 2006–2007
Jonathan D. Klein, MD, MPH, Chairperson
Michelle S. Barratt, MD, MPH
*Margaret J. Blythe, MD
Paula K. Braverman, MD
*Angela Diaz, MD
David S. Rosen, MD, MPH
Charles J. Wibbelsman, MD

PAST COMMITTEE MEMBERS
Ronald A. Feinstein, MD
Martin M. Fisher, MD
David W. Kaplan, MD, MPH
Ellen S. Rome, MD, MPH
W. Samuel Yancy, MD

LIAISONS
Miriam Kaufman, RN, MD
Canadian Paediatric Society
Lesley Breech, MD
American College of Obstetricians and Gynecologists
Benjamin Shain, MD, PhD
American Academy of Child and Adolescent Psychiatry
S. Paige Hertweck, MD
American College of Obstetricians and Gynecologists

STAFF
Karen Smith
*Lead authors

REFERENCES

1. Haignere CS, Gold R, McDanel HJ. Adolescent abstinence and condom use: are we sure we are really teaching what is safe? Health Educ Behav. 1999;26:43–54


95. American Academy of Pediatrics, Committee on Adolescence. 
104. Jackson RA, Biimla Schwarz E, Freedman L, Darney P. Advance supply of emergency contraception: effect on use and usual contraception—a randomized trial. *Obstet Gynecol*. 2003;102:8–16

**ADDITIONAL RESOURCES**
Miller LJ. Sexuality, reproduction, and family planning in women with schizophrenia. *Schizophr Bull*. 1997;23:623–635
Global Climate Change and Children’s Health

Committee on Environmental Health

ABSTRACT
There is broad scientific consensus that Earth’s climate is warming rapidly and at an accelerating rate. Human activities, primarily the burning of fossil fuels, are very likely (>90% probability) to be the main cause of this warming. Climate-sensitive changes in ecosystems are already being observed, and fundamental, potentially irreversible, ecological changes may occur in the coming decades. Conservative environmental estimates of the impact of climate changes that are already in process indicate that they will result in numerous health effects to children. The nature and extent of these changes will be greatly affected by actions taken or not taken now at the global level.

Physicians have written on the projected effects of climate change on public health, but little has been written specifically on anticipated effects of climate change on children’s health. Children represent a particularly vulnerable group that is likely to suffer disproportionately from both direct and indirect adverse health effects of climate change. Pediatric health care professionals should understand these threats, anticipate their effects on children’s health, and participate as children’s advocates for strong mitigation and adaptation strategies now. Any solutions that address climate change must be developed within the context of overall sustainability (the use of resources by the current generation to meet current needs while ensuring that future generations will be able to meet their needs). Pediatric health care professionals can be leaders in a move away from a traditional focus on disease prevention to a broad, integrated focus on sustainability as synonymous with health.

This policy statement is supported by a technical report that examines in some depth the nature of the problem of climate change, likely effects on children’s health as a result of climate change, and the critical importance of responding promptly and aggressively to reduce activities that are contributing to this change.

BACKGROUND
“Warming of the climate system is unequivocal.” According to the National Climatic Data Center, all records indicate that during the past century, global surface temperatures have increased at a rate near 0.6°C per century (1.1°F per century); this trend has been 3 times larger since 1976. Human activity, particularly the burning of fossil fuels, has very likely (>90% probability) driven this rise by greatly increasing atmospheric concentrations of carbon dioxide (CO2) and other greenhouse gases (GHGs). There is strong consensus among expert scientists that Earth is undergoing rapid, global climate change, although there remains uncertainty about how rapidly and extensively the climate will change in the future. Overall scientific predictions agree, however, that temperatures and sea level will continue to rise.
throughout the 21st century.\textsuperscript{1,4} Even if GHG emissions were abruptly reduced to zero, the planet would continue to warm for decades until the energy stored in the system equilibrates.\textsuperscript{3} The possibility of reaching a tipping point at which abrupt, large, and irreversible change could be superimposed on current trends adds both urgency and further ambiguity to the situation.\textsuperscript{6} Current human activities are accelerating these changes, and future human activities will affect their trajectories; the window of opportunity for successful mitigation, therefore, may be very short.\textsuperscript{7} Actions made in the coming decade will have a profound effect on global health and, in particular, on children’s health.

**DIRECT EFFECTS OF CLIMATE CHANGE ON CHILDREN’S HEALTH**

Because of their physical, physiologic, and cognitive immaturity, children are often most vulnerable to adverse health effects from environmental hazards.\textsuperscript{8} As the climate changes, environmental hazards may shift and possibly increase (Fig 1), and children are likely to suffer disproportionately from these changes.\textsuperscript{9} Anticipated direct health consequences of climate change include injury and death from extreme weather events and natural disasters, increases in climate-sensitive infectious diseases, increases in air pollution–related illness, and more heat-related, potentially fatal, illness. Within all of these categories, children have increased vulnerability compared with other groups (see the accompanying technical report\textsuperscript{10}).

**INDIRECT EFFECTS OF CLIMATE CHANGE AND IMPLICATIONS FOR FUTURE GENERATIONS**

Additional effects of climate change, with profound implications for the health and welfare of future generations of children, are anticipated. Food availability could be reduced as land and ocean food productivity patterns shift and species diversity declines.\textsuperscript{11} Water availability will change and become too abundant in some regions (flooding) and much reduced in others (drought).\textsuperscript{12} Coastal populations will be forced to move because of the rising sea level. Large-scale, forced migrations are conceivable, driven by abrupt climate change, natural disaster, or political instability over resource availability.\textsuperscript{13}

The speed with which global GHG emissions can be reduced will have a significant effect on the rate and degree of warming, but even the most optimistic scenarios describe continued warming into the next century.\textsuperscript{1,5} As climate change progresses, social and political institutions must respond with aggressive mitigation and flexible adaptation strategies to preserve and protect public health, particularly for children.

---

**FIGURE 1**

Potential effects of global climate change on child health. (Adapted from McMichael et al\textsuperscript{18} and Haines and Patz\textsuperscript{19}.)
MITIGATION AND ADAPTATION STRATEGIES

Strategies to address the effects of climate change (mitigation and adaptation) are concepts that focus on both primary and secondary prevention strategies in pediatric health care (Fig 1). Mitigation (primary prevention) involves reducing GHG concentrations in the atmosphere with the goal of reducing climate change. Adaptation (secondary prevention) involves developing public health strategies to minimize and, in some cases, eliminate local and regional adverse health outcomes that are anticipated from climate change.

A wide variety of governmental and nongovernmental organizations have developed detailed lists of mitigation and adaptation strategies, from international conventions such as the Kyoto Protocol to individual actions such as reducing automobile use.

However, any solutions that address climate change must be developed within the context of overall sustainable development. The use of resources by the current generation to meet current needs while ensuring that future generations will be able to meet their needs.

Given the health implications of climate change for current and future generations of children, the disease-prevention role for pediatric health care professionals includes advocating for environmental sustainability.

RECOMMENDATIONS TO PEDIATRICIANS

Pediatricians are dedicated to the promotion and protection of children’s health. Climate change threatens the health, welfare, and future of current and subsequent generations of children. Pediatricians can incorporate considerations of the effects of climate change on health into their professional practice and personal lives in many ways, including patient education, lifestyle practices, and political advocacy. Some possible approaches might include the following.

1. Recognize and educate yourself about the links between child health and climate change. Existing anticipatory guidance already incorporates many issues that can help mitigate climate change. For example, encouraging families and children to walk or ride bicycles more may reduce automobile emissions.

2. Advocate for comprehensive local and national policies that address climate change to improve the health of children now and in the future. Educate elected officials on the health risks to children from climate change; write letters to the editor, attend public meetings, or provide expert testimony. Work with local schools, child care centers, community organizations, and businesses on projects that will help reduce GHGs. Support policies to expand parks and green spaces, strengthen public transport, improve sidewalks and bicycle lanes, and create local award systems for energy-efficient businesses, buildings, organizations, and households.

3. Serve as a role model for practices that promote environmental sustainability. Emphasize energy conservation in your workplace, encourage and model reduced dependency on automobile travel, and consider the environmental and energy costs when making major purchases for your practice or institution.

4. Help to build and support coalitions across disciplines and institutions to search for novel, comprehensive approaches to mitigate and adapt to climate change in your community and region. Work with local and state health departments to strengthen public health infrastructure, disease surveillance and reporting, and disaster preparedness.

5. Work to ensure that concepts related to the pediatric health implications of climate change are part of pediatric training and curricula.

RECOMMENDATIONS TO GOVERNMENT

Government at all levels, from the smallest municipalities to the national and international levels, should implement aggressive policies to halt man-made contributions to climate change and to mitigate its impact on children’s health.

1. Develop aggressive, long-term policies to reduce the major contributing factors to global climate change.

2. Invest in prudent and vital preparations for our public health care systems, including immunization programs and disease surveillance, reporting, and tracking.

3. Give specific attention to the needs of children in emergency management and disaster response.

4. Support education and public awareness of the threats from climate change and their implications for public and children’s health now and in the future.

5. Fund interdisciplinary research to develop, implement, and measure outcomes of innovative strategies to both mitigate and adapt to climate change, particularly in areas with direct implications for children’s health.

COMMITTEE ON ENVIRONMENTAL HEALTH, 2006–2007

Michael W. Shannon, MD, MPH, Chairperson
Dana Best, MD, MPH
Helen J. Binns, MD, MPH
Joel A. Forman, MD
Christine L. Johnson, MD
Catherine J. Karr, MD, PhD
Janice J. Kim, MD, PhD, MPH
Lynnette J. Mazur, MD, MPH
James R. Roberts, MD, MPH

PAST COMMITTEE MEMBER
*Katherine M. Shea, MD, MPH

LIAISONS

Elizabeth Blackburn, RN
US Environmental Protection Agency
REFERENCES
Role of the Medical Home in Family-Centered Early Intervention Services

Council on Children With Disabilities

ABSTRACT
There is growing evidence that early intervention services have a positive influence on the developmental outcome of children with established disabilities as well as those who are considered to be “at risk” of disabilities. Various federal and state laws now mandate the establishment of community-based, coordinated, multidisciplinary, family-centered programs that are accessible to children and families. The medical home, in close collaboration with the family and the early intervention team, can play a critical role in ensuring that at-risk children receive appropriate clinical and developmental early intervention services. The purpose of this statement is to assist the pediatric health care professional in assuming a proactive role with the interdisciplinary team that provides early intervention services.

EARLY INTERVENTION LEGISLATION
Various federal and state laws now mandate the establishment of community-based, coordinated, multidisciplinary, family-centered programs that are accessible to children with established disabilities or those who are “at risk” of disabilities and their families. Early intervention services are designed to meet the needs of children from birth to 36 months of age who have delays in 1 or more areas of physical, cognitive, communication, social, emotional, or adaptive development. Services are also available to children who have a diagnosed condition that has a high probability of resulting in delayed development. States must offer early intervention services to children with delayed development or those with an established disability. States also have the option of serving those who are at risk for poor developmental outcomes. The type and extent of services are determined through the development of an individualized family service plan (IFSP). In designing the IFSP, the family plays a lead role in the assessment of resources, priorities, and concerns in conjunction with a care coordinator.1,2

By federal statute, available services must include:

- early identification, screening, and assessment services;
- care-coordination services;
- medical services only for diagnostic or evaluation purposes;
- family training, counseling, and home visits;
- special instruction;
- speech and language pathology and audiology services;
- occupational and physical therapy;

Key Words
early intervention

Abbreviations
IFSP—individualized family service plan
IDEA—Individuals With Disabilities Education Act
AAP—American Academy of Pediatrics

www.pediatrics.org/cgi/doi/10.1542/peds.2007-2638
doi:10.1542/peds.2007-2638
All policy statements from the American Academy of Pediatrics automatically expire 5 years after publication unless reaffirmed, revised, or retired at or before that time.

PEDIATRICS Volume 120, Number 5, November 2007 1153
the states that did not serve the at-risk population to track and monitor these children so that they could be referred when needed.3,4

The Individuals With Disabilities Education Improvement Act of 2004 (IDEA 2004 [Pub L No. 108-446]) broadened the eligibility criteria for early intervention services. The 2004 legislation required referral for all children involved in substantiated cases of neglect or abuse, children affected by substance abuse or exposed to family violence, and children who are homeless or wards of the state. IDEA 2004 also permitted, at the states’ discretion, families to choose to have their child continue in early intervention services until they are eligible for kindergarten.5

RATIONALE FOR EARLY INTERVENTION
Until 3 decades ago, in the absence of laws that mandated access to educational services for all children regardless of the degree of disability, many children with developmental disabilities and their families had few choices except state hospital-sponsored custodial care or an isolated homebound existence. Since then, much has been accomplished in the field of health care and special education for children with disabilities. Recent advances in medical expertise and technology have improved the developmental potential, health, and survival rate of infants and children with special health care needs. These advances have enabled children with special health care needs to participate more fully in public education. Neurocognitive research has demonstrated that there are optimal periods for all children during which the brain is particularly efficient at specific types of learning. Well-designed, timely early intervention can improve the outcome and the quality of life of young children at risk of developing cognitive, social, or emotional impairment.6–9 The early childhood years present a singular opportunity to influence lifelong development and prevent or minimize developmental problems in children with disabilities or those who are at risk of developing disabilities.

THE BENEFITS OF EARLY INTERVENTION
Pediatric health care professionals have a major role in early identification and referral for children with established delays in development as well as children who are at risk of delays. The National Early Intervention Longitudinal Study10 found that the age at first concerns was later for children with developmental delays (11.1 months) compared with children with diagnosed conditions (eg, Down syndrome) (2.3 months) and children with at-risk conditions (eg, prematurity) (2.1 months). The time between first concerns and development of an IFSP was also longer for children with developmental delays (8.9 months) compared with children with diagnosed conditions (7.1 months) and children with risk conditions (5.9 months). Children with developmental

- psychological services;
- health services that are necessary to enable the infant or toddler to benefit from other early intervention services;
- social work services;
- vision services;
- assistive technology devices and services; and
- transportation, interpretation services, and other related costs that are necessary to enable a family to receive other services.3,4

Access to these services has been mandated because early intervention is important if children with disabilities are to achieve their full potential. During the past 25 years, the US Congress has taken a series of steps to promote improved infant and child developmental outcomes through early intervention services. The first major federal legislation was passed in 1975, when the Education for All Handicapped Children Act (Pub L No. 94-142) established the right of children between 5 and 18 years of age to a free, appropriate public education and related services; providing services to children 3 to 5 years of age was optional. The Education of the Handicapped Amendments of 1986 (Pub L No. 99-457) supported the development of early intervention programs for infants and children from birth to 3 years of age with disabilities or delayed development. The law also mandated that a free and appropriate public education be provided by the states’ education departments for 3- to 5-year-olds by the 1990–1991 school year. It established guidelines and regulations for the development of far-reaching, coordinated, multidisciplinary services for these children and their families. In 1990, it was amended again as the Individuals With Disabilities Education Act (IDEA [Pub L No. 101-476]). One component of IDEA, Part H (now known as Part C), the Program for Infants and Toddlers With Disabilities, required states to develop and implement community-based systems of care that are coordinated, family centered, and culturally effective, with greater interagency collaboration. Part H required early identification and provision of services to infants and toddlers with delayed development and those with established conditions with a high probability of delay and, at the state’s option, those who would be at risk of experiencing delayed development if early intervention services were not provided. Part H required that identified children be referred for a free comprehensive, multidisciplinary evaluation by a team of professionals who, with the family, decide which services are needed. The services that are determined to be necessary are listed on the IFSP, and the needs are reevaluated at least annually. A care coordinator is appointed to help the family access services. Subsequently, Part C of the IDEA Amendments of 1997 (Pub L No. 105-17) encouraged
delays were older than children with diagnosed conditions and risk conditions at the time of the IFSP. Male children with delays entered services at later ages than did female children with delays. White children with delays entered services slightly later than did children of other ethnicities with delays. No gender or ethnicity differences regarding age at entry within diagnosed conditions or at-risk groups were found. Sixty-four percent of families found doctors or other health professionals to be very helpful. Most parents felt that early intervention services helped their child’s development and that their family was better off with these services. These findings were not as strong for low-income families or if the child had poor health.

These data suggest that pediatric health care professionals can improve early identification and referral for children at biological and environmental risks as well as those with delayed development without known risk factors. The American Academy of Pediatrics (AAP) has published an algorithm for developmental surveillance and screening in early childhood that can assist the medical home in this process. The Infant Health and Development Program is a multicentered, randomized, controlled, nationwide study of low birth weight preterm infants (and their families) who received coordinated health and developmental services for the first 3 years of life. Children who had received comprehensive, multidisciplinary early intervention services scored higher at 3 years of age on tests of mental abilities than did children who received health services alone. Within the intervention group, cognitive and academic achievement in children with higher birth weight was maintained at 8 years of age. School outcomes for children in the intervention group were consistently better than for children who did not receive intervention. Several aspects of family development were also enhanced by the Infant Health and Development Program.

Another long-term study, the Carolina Abecedarian Project, recently revealed that poor children who received early educational intervention starting in infancy had higher scores on mental, reading, and math tests than did children who did not receive the intervention. The participants were assessed at 21 years of age and were found to have completed more years of education, were more likely to attend a 4-year college, and were older when their first child was born.23

There has been considerable growth in the field of research regarding efficacy of various treatment modalities for children with specific disabilities. It is important to consider this research when prescribing or providing advice regarding early intervention services. For example, for those with cerebral palsy, data suggest that a functional/behavioral approach warrants initial consideration. Muscle strength training should also be considered for children with cerebral palsy. Additional guidelines for prescribing therapy services for children with motor disabilities were published by the AAP in 2004.24

Lipkin and Schertz’s review of the literature on early intervention for children with Down syndrome suggested that early intervention may be beneficial in preventing declines in IQ. Preliminary findings have raised promise for treadmill training and augmentative communication to improve outcomes.

Evidence for the benefits of early intervention for children with autism is stronger. The evidence suggests that early, intensive (at least 20 hours/week) behavioral and/or developmental services are helpful in improving communication and social skills, but more research is needed (including ongoing research) regarding the types and intensity of services.

The parents and family, as the primary caregivers, play a vital role in ensuring the health and well-being of children. The focus of health and developmental services has evolved from a child-centered, traditional “medical” model to a family-centered “developmental” model. That is, those who coordinate services take into consideration the important contributions of the family unit, the stressors that affect families (social, financial, and/or psychological), and the ability of families to adapt to new challenges. The pediatric health care professional, as the central figure in the medical home, must be attuned to special family circumstances that influence children with
special health care needs. The pediatric health care professional must involve family members in all areas of planning, delivery, and evaluation of health and developmental services. Communication between parents and pediatric health care professionals should be open, comprehensible, culturally sensitive, and sincere, showing mutual respect.

The pediatric health care professional, because of his or her unique training, interest, and commitment, should be a vital member of the early intervention health team. The pediatric health care professional is the most appropriate health care consultant, coordinator, and source of referral for clinical services for children with special health care needs and their families. Whether in a local pediatric health care professional’s office or in a multispecialty referral center, these children and their families should be offered comprehensive care that is family centered, continuous, compassionate, and culturally sensitive. Regardless of the pediatric health care setting, this care can be provided in accordance with the precepts of the medical home.

RECOMMENDATIONS

The role of the pediatric health care professional caring for children with disabilities and their families should include:

- Surveillance and screening of all infants to identify established disabilities or risks of delayed development following the AAP algorithm. The algorithm contains recommendations to perform surveillance at all well-child visits and administration of a standardized screening tool at the 9- and 18-month visits and again at either the 24- or 30-month visit.
- Referring children with delayed development or established risk factors promptly to early intervention services. The AAP and the US Department of Education Office of Special Education Programs have collaborated to develop a referral form, which accompanies this statement.
- Arranging for medical etiologic diagnostic evaluation as appropriate. Guidelines for evaluation of children with delayed development have been published by the AAP and the American Academy of Neurology. Guidelines for diagnostic assessment of cerebral palsy also are available. In addition, the AAP, the American Academy of Neurology, and the American Academy of Child and Adolescent Psychiatry have published guidelines for assessment of children with autistic spectrum disorders.
- Being aware of the services and resources available in the community for the child and family and helping to coordinate the health component of the services.
- Collaborating with the family and care coordinator to provide medical input into development of the IFSP while ensuring that goals are functional in nature. Efforts at collaboration have been hampered by lack of payment for these services.
- Advocating for the child’s access to the appropriate medical subspecialty and surgical specialty services.
- Supporting families in choosing evidence-based and best practices that meet the specific needs of their child.
- Ensuring that periodic, objective measures of progress are made and used to guide ongoing intervention design.
- Providing continuity of health care, including prescribing specific rehabilitative therapies as appropriate and periodically reviewing the need to continue such services on the basis of the achievement of common goals.
- Periodic and ongoing counseling for the family regarding the child’s progress and treatment and management options.
- Helping to provide ongoing services that are aimed at preventing secondary disabilities.
- Maintaining a central medical database that contains pertinent diagnostic and consultative information.
- Negotiating for proper payment for time and effort spent on care coordination, counseling services, and other direct services.
- Advocating for equal access to early intervention programs for all eligible children in need.
- Advocating for ongoing evaluation of early intervention programs through quality assurance and other performance measures.
- Representing state AAP chapters on local and state interagency coordination councils.
- Monitoring and supporting research that uses optimal methodologies to further clarify appropriate treatment modalities for children with specific disabilities.

CONCLUSIONS

By providing leadership for the medical home and as a member of the early intervention team, pediatric health care professionals can help set the standard of care in their communities for children with disabilities or those who are at risk of developmental delays. Through ongoing consultation with developmental and rehabilitation therapists, services and therapy prescriptions should be provided with specific treatment goals in mind. Treatment plans should be regularly and periodically reviewed and revised, if necessary, or renewed if indications show that they are accomplishing their intended purpose.

It is vital for pediatric health care professionals to be sensitive to their role as the medical care provider on the
early intervention team, promoting appropriate education and therapy for children with disabilities. An environment should be created in which the pediatric health care professional, family, and other service providers work together in a caring, collegial, and compassionate atmosphere that ensures that early intervention services are of high quality, accessible, continuous, comprehensive, and culturally effective.

COUNCIL ON CHILDREN WITH DISABILITIES, 2006–2007
Paul H. Lipkin, MD, Chairperson
James Daniel Cartwright, MD
Larry W. Desch, MD
*John C. Duby, MD
Ellen Roy Elias, MD
Chris Plauché Johnson, MD, MEd
Eric B. Levey, MD
Gregory S. Liptak, MD
Larry W. Desch, MD
James Daniel Cartwright, MD
Paul H. Lipkin, MD, Chairperson

REFERENCES

5. Individuals With Disabilities Education Improvement Act of 2004 (Pub L No. 108-446)
26. Ramey CT, Ramey SL. Early intervention: optimizing develop-


Ventricular Fibrillation and the Use of Automated External Defibrillators on Children

Committee on Pediatric Emergency Medicine and Section on Cardiology and Cardiac Surgery

ABSTRACT
The use of automated external defibrillators (AEDs) has been advocated in recent years as one part of the chain of survival to improve outcomes for adult cardiac arrest victims. When AEDs first entered the market, they had not been tested for pediatric usage and rhythm interpretation. In addition, the presumption was that children do not experience ventricular fibrillation, so they would not benefit from the use of AEDs. Recent literature has shown that children do experience ventricular fibrillation, which has a better outcome than do other cardiac arrest rhythms. At the same time, the arrhythmia software on AEDs has become more extensive and validated for children, and attenuation devices have become available to downregulate the energy delivered by AEDs to allow their use on children. Pediatricians are now being asked whether AED programs should be implemented, and where they are being implemented, pediatricians are being asked to provide guidance on the use of them on children. As AED programs expand, pediatricians must advocate on behalf of children so that their needs are accounted for. For pediatricians to be able to provide guidance and ensure that children are included in AED programs, it is important for pediatricians to know how AEDs work, be up-to-date on the literature regarding pediatric fibrillation and energy delivery, and understand the role of AEDs as life-saving interventions for children.

INTRODUCTION
Early defibrillation has been shown to be the most effective treatment for adult out-of-hospital cardiac arrest attributable to ventricular fibrillation (VF).1,2 The likelihood of survival decreases by approximately 7% to 10% with each minute of delay to defibrillation after cardiac arrest. Strategies to decrease the time to defibrillation that have been shown to be effective include the use of an automated external defibrillator (AED) by prehospital care providers and nonmedical personnel.3-6

For children, use of defibrillation traditionally has been downplayed, with a focus on early airway and ventilatory assistance as a result of data that showed that asystole was the predominant rhythm and that VF rarely occurred.7 Although not the most common rhythm, VF does occur in children. In addition, the chance of survival after VF is greater than that from other nonperfusing rhythms, which makes treatment of VF a priority in pediatric resuscitation.8

Although the incidence of VF in the pediatric population is low, there is a need for developing strategies to provide early defibrillation to patients younger than 8 years. This strategy may include the need for an AED that is suitable for use in pediatric
patients from birth to 8 years of age with either an attenuated adult-dosage AED tested for efficacy and safety in children or an AED specifically designed for defibrillation of young children and infants. In the absence of either of such devices, standard nonattenuated adult-dosage AEDs should be used on children on the basis of protocols for use developed with medical oversight. Because of the limited data on effective energy dose, emergency medical services (EMS) systems, medical directors, and pediatric researchers should make efforts to gather information regarding pediatric uses of AEDs and report it by using the pediatric Utstein style, which represents an internationally accepted standard method of collecting and reporting respiratory and cardiac arrest and resuscitation data. In addition, because the use of AEDs on children may be a new concept to many responders, EMS and physician leaders should work with professional organizations, community organizations, and researchers to educate their community members regarding the benefits of early pediatric defibrillation and the use of available varieties of AEDs.

The message for the public and for EMS systems is to recognize the existence of VF in infants and children and use methods to treat it as early as possible to improve the survival of children and infants after sudden cardiac arrest. In addition, this possible life-saving therapy should not be withheld purely on the basis of absolute weight and size issues. The key is to pursue a long-term goal of providing devices that will support rapid pediatric and adult defibrillation. This approach would include the ability to treat infants and children without compromising adult care while minimizing training issues and minimizing the use of limited financial and personnel resources.

RECOMMENDATIONS

1. Although the incidence of VF in children is far less than that in adults, the outcome for VF is better than that for other nonperfusing rhythms and is improved with early defibrillation. Therefore, strategies and equipment availability for treatment of pediatric arrest should be focused on shortening the intervals from collapse to recognition of VF and to defibrillation.

2. Although most data available on the correct energy for defibrillation of children are from animal studies, data suggest that the immature heart is less susceptible to energy-related damage and that there is a wide therapeutic range of defibrillation energy doses. In addition, although using a fixed adult-energy AED on some children has a theoretical potential for harm, not treating VF has the proven potential for even greater harm: death of the child. On the basis of this risk/benefit assessment, prehospital programs and public-access AED programs should not withhold defibrillation because of weight or age criteria alone.

3. Children and infants of all ages who suffer VF must be provided defibrillation as soon as possible after arrest. The following approach to achieve this goal should be used:
   a. Immediately provide defibrillation to all infants and children from birth to 8 years of age with either an attenuated adult-dosage AED tested for efficacy and safety in children or an AED specifically designed for defibrillation of young children and infants, depending on which device is available first. In the absence of the devices listed above, standard nonattenuated adult-dosage AEDs should be used on infants and children from birth to 8 years of age. Protocols for use of adult-dosage AEDs should be developed with medical oversight.
   b. For children 8 years of age and older, immediately provide defibrillation with an adult-dosage AED or manual defibrillation.

4. EMS systems must have protocols to allow for pediatric defibrillation in the timeliest fashion and by all levels of responders. These protocols include pediatric AED capability and, in the interim, protocols for the use of an adult AED on infants and children.

5. Although a cost/benefit assessment of public-access defibrillation specifically for children has not been established, when a community or facility chooses to establish a public-access defibrillation program, the AED chosen for that program must have pediatric capability.

6. Although a cost/benefit assessment of school-based AEDs has not been established yet, school systems must, in their assessment of need for an AED, consider the benefit of AED purchase to adult staff members and adult visitors and as another component of school-based emergency care.

7. When determining the need for a school-based AED program, the following factors should direct the decision:
   a. The frequency of cardiac arrest events is such that there is a reasonable probability of AED use within 5 years of rescuer training and AED placement. This probability can be established if 1 cardiac arrest has been known to have occurred at the site within the last 5 years or can be estimated on the basis of population demographics.
   b. There are children attending the school or adults working at the school who are thought to be at high risk of sudden cardiac arrest (eg, children with conditions such as congenital heart disease and a history of abnormal heart rhythms, children with long QT syndrome, children with cardiomyopathy, adults or children who have had a heart transplant, and adults with a history of heart disease).
c. An EMS call-to-shock interval of less than 5 minutes cannot be achieved reliably with conventional EMS services, and a collapse-to-shock interval of less than 5 minutes can be achieved reliably (in >90% of cases) by training and equipping lay people to function as first responders by recognizing cardiac arrest, telephoning 911 (or other appropriate emergency response number), starting cardiopulmonary resuscitation, and attaching and operating an AED.

8. When placed in schools, AEDs must be part of a comprehensive emergency care plan that includes:
   a. pediatric medical oversight;
   b. staff training in basic first aid and cardiopulmonary resuscitation; and
   c. integration with local EMS.

9. Any legislation that mandates placement of an AED also must provide the funding for such devices, including costs of staff training and maintenance of the equipment.

10. AED legislation must allow for pediatric AED usage and liability protection for those who use these devices and for the physicians who provide the medical oversight for these programs.

COMMITTEE ON PEDIATRIC EMERGENCY MEDICINE, 2005–2006
Steven E. Krug, MD, Chairperson
Thomas Bojko, MD, MS
Margaret A. Dolan, MD
Karen S. Frush, MD
Patricia J. O’Malley, MD
Robert E. Sapien, MD
Kathy N. Shaw, MD, MSCE
Joan Shook, MD, MBA
Paul E. Sirbaugh, DO
Loren G. Yamamoto, MD, MPH, MBA

LIAISONS
Jane Ball, RN, DrPH
   EMSC National Resource Center
Kathleen Brown, MD
   National Association of EMS Physicians
Kim Bullock, MD
   American Academy of Family Physicians
Dan Kavanaugh, MSW
   Maternal and Child Health Bureau
Sharon E. Mace, MD
   American College of Emergency Physicians
Susan Eads Role, JD, MSLS
   EMSC National Resource Center
David W. Tuggle, MD
   American College of Surgeons
Tina Turgel, BSN, RN-C
   Maternal and Child Health Bureau

CONTRIBUTOR
David Markenson, MD

STAFF
Susan Tellez

SECTION ON CARDIOLOGY AND CARDIAC SURGERY, 2006–2007
Robert H. Beekman III, MD, Chairperson
Peter B. Manning, MD
Seema Mital, MD
William R. Morrow, MD
Frank M. Galioto, Jr, MD
Thomas K. Jones, MD

LIAISONS
Gerard R. Martin, MD
American College of Cardiology
Reginald L. Washington, MD
NCE Planning Group

STAFF
Lynn Colegrove, MBA

REFERENCES
CLINICAL REPORT

Management of Children With Autism Spectrum Disorders

Scott M. Myers, MD, Chris Plauche Johnson, MD, MEd, the Council on Children With Disabilities

ABSTRACT

Pediatricians have an important role not only in early recognition and evaluation of autism spectrum disorders but also in chronic management of these disorders. The primary goals of treatment are to maximize the child’s ultimate functional independence and quality of life by minimizing the core autism spectrum disorder features, facilitating development and learning, promoting socialization, reducing maladaptive behaviors, and educating and supporting families. To assist pediatricians in educating families and guiding them toward empirically supported interventions for their children, this report reviews the educational strategies and associated therapies that are the primary treatments for children with autism spectrum disorders. Optimization of health care is likely to have a positive effect on habilitative progress, functional outcome, and quality of life; therefore, important issues, such as management of associated medical problems, pharmacologic and nonpharmacologic intervention for challenging behaviors or coexisting mental health conditions, and use of complementary and alternative medical treatments, are also addressed.

INTRODUCTION

The term autism spectrum disorders (ASDs) has been used to include the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) 1 diagnostic categories autistic disorder, Asperger disorder, and pervasive developmental disorder—not otherwise specified. 2 Recent estimates of the prevalence of ASDs are in the range of 6.5 to 6.6 per 1000, and pediatricians, therefore, are likely to care for children and adolescents with these diagnoses. 3-5 In the companion document to this clinical report, 2 the American Academy of Pediatrics has summarized pertinent background information on ASDs and emphasized the importance of surveillance and screening as well as other potential physician roles in the diagnostic process. However, the role of the primary health care professional extends beyond recognizing signs of ASDs, referring for diagnostic evaluation, conducting an etiologic investigation, providing genetic counseling, and educating caregivers about ASDs and includes ongoing care and management.

ASDs, similar to other neurodevelopmental disabilities, are generally not “curable,” and chronic management is required. Although outcomes are variable and specific behavioral characteristics change over time, most children with ASDs remain within the spectrum as adults and, regardless of their intellectual functioning, continue to experience problems with independent living, employment, social relationships, and mental health. 6-8 The primary goals of treatment are to minimize the core features and associated deficits, maximize functional indepen-
dence and quality of life, and alleviate family distress. Facilitating development and learning, promoting socialization, reducing maladaptive behaviors, and educating and supporting families can help accomplish these goals. Ideally, interventions should help mitigate the core features of ASDs, which include impairment in social reciprocity, deficits in communication, and restricted, repetitive behavioral repertoire.

Educational interventions, including behavioral strategies and habilitative therapies, are the cornerstones of management of ASDs. These interventions address communication, social skills, daily-living skills, play and leisure skills, academic achievement, and maladaptive behaviors.

Optimization of medical care is also likely to have a positive impact on habilitative progress and quality of life. In addition to routine preventive care and treatment of acute illnesses, management of sleep dysfunction, coexisting challenging behaviors or psychiatric conditions, and associated medical problems, such as seizures, may be particularly important. Medications have not been proven to correct the core deficits of ASDs and are not the primary treatment. However, associated maladaptive behaviors or psychiatric comorbidities may interfere with educational progress, socialization, health or safety, and quality of life. These behaviors may be amenable to psychopharmacologic intervention or, in some cases, treatment of underlying medical conditions that are causing or exacerbating the behaviors. Effective medical management may allow a child with an ASD to benefit more optimally from educational interventions.

**EDUCATIONAL INTERVENTIONS**

Education has been defined as the fostering of acquisition of skills and knowledge to assist a child to develop independence and personal responsibility; it encompasses not only academic learning but also socialization, adaptive skills, communication, amelioration of interfering behaviors, and generalization of abilities across multiple environments. Physicians and other clinicians are often in a position to guide families to empirically supported practices and help them evaluate the appropriateness of the educational services that are being offered.

**Comprehensive Programs for Young Children**

In the last 2 decades, research and program development in the area of educational intervention have focused largely on very young children with ASDs because of earlier identification and evidence that early intensive intervention may result in substantially better outcomes. Model early childhood educational programs for children with ASDs have been described in thorough reviews. These model programs are often categorized as behavior analytic, developmental, or structured teaching on the basis of the primary philosophical orientation. Although the approaches have important differences, they also overlap. For example, contemporary comprehensive behavioral curricula borrow from developmental or cognitive approaches (such as addressing joint attention, reciprocal imitation, symbolic play, and theory of mind and using indirect language stimulation and contingent imitation techniques), and some developmental models (eg, the Denver model) and the structured teaching approach of the Treatment and Education of Autistic and Related Communication Handicapped Children (TEACCH) program use behavioral techniques to fulfill their curriculum goals.

Although programs may differ in philosophy and relative emphasis on particular strategies, they share many common goals, and there is a growing consensus that important principles and components of effective early childhood intervention for children with ASDs include the following:

- entry into intervention as soon as an ASD diagnosis is seriously considered rather than deferring until a definitive diagnosis is made;
- provision of intensive intervention, with active engagement of the child at least 25 hours per week, 12 months per year, in systematically planned, developmentally appropriate educational activities designed to address identified objectives;
- low student-to-teacher ratio to allow sufficient amounts of 1-on-1 time and small-group instruction to meet specific individualized goals;
- inclusion of a family component (including parent training as indicated);
- promotion of opportunities for interaction with typically developing peers to the extent that these opportunities are helpful in addressing specified educational goals;
- ongoing measurement and documentation of the individual child’s progress toward educational objectives, resulting in adjustments in programming when indicated;
- incorporation of a high degree of structure through elements such as predictable routine, visual activity schedules, and clear physical boundaries to minimize distractions;
- implementation of strategies to apply learned skills to new environments and situations (generalization) and to maintain functional use of these skills; and
- use of assessment-based curricula that address:
  - functional, spontaneous communication;
  - social skills, including joint attention, imitation, reciprocal interaction, initiation, and self-management;
• functional adaptive skills that prepare the child for increased responsibility and independence;
• reduction of disruptive or maladaptive behavior by using empirically supported strategies, including functional assessment;
• cognitive skills, such as symbolic play and perspective taking; and
• traditional readiness skills and academic skills as developmentally indicated.

Specific Strategies
A variety of specific methodologies are used in educational programs for children with ASDs. Detailed reviews of intervention strategies to enhance communication,
• traditional readiness skills and academic skills as developmentally indicated.

Specific Strategies
A variety of specific methodologies are used in educational programs for children with ASDs. Detailed reviews of intervention strategies to enhance communication, teach social skills, and reduce interfering maladaptive behaviors have been published in recent years. Brief descriptions of selected methodologies are provided below.

Applied Behavior Analysis
Applied behavior analysis (ABA) is the process of applying interventions that are based on the principles of learning derived from experimental psychology research to systematically change behavior and to demonstrate that the interventions used are responsible for the observable improvement in behavior. ABA methods are used to increase and maintain desirable adaptive behaviors, reduce interfering maladaptive behaviors or narrow the conditions under which they occur, teach new skills, and generalize behaviors to new environments or situations. ABA focuses on the reliable measurement and objective evaluation of observable behavior within relevant settings including the home, school, and community. The effectiveness of ABA-based intervention in ASDs has been well documented through 5 decades of research by using single-subject methodology and in controlled studies of comprehensive early intensive behavioral intervention programs in university and community settings. Children who receive early intensive behavioral treatment have been shown to make substantial, sustained gains in IQ, language, academic performance, and adaptive behavior as well as some measures of social behavior, and their outcomes have been significantly better than those of children in control groups.

Highly structured comprehensive early intervention programs for children with ASDs, such as the Young Autism Project developed by Lovaas at the University of California Los Angeles, rely heavily on discrete trial training (DTT) methodology, but this is only one of many techniques used within the realm of ABA. DTT methods are useful in establishing learning readiness by teaching foundation skills such as attention, compliance, imitation, and discrimination learning, as well as a variety of other skills. However, DTT has been criticized because of problems with generalization of learned behaviors to spontaneous use in natural environments and because the highly structured teaching environment is not representative of natural adult-child interactions. Traditional ABA techniques have been modified to address these issues. Naturalistic behavioral interventions, such as incidental teaching and natural language paradigm/pivotal response training, may enhance generalization of skills.

Functional behavior analysis, or functional assessment, is an important aspect of behaviorally based treatment of unwanted behaviors. Most problem behaviors serve an adaptive function of some type and are reinforced by their consequences, such as attainment of (1) adult attention, (2) a desired object, activity, or sensation, or (3) escape from an undesired situation or demand. Functional assessment is a rigorous, empirically based method of gathering information that can be used to maximize the effectiveness and efficiency of behavioral support interventions. It includes formulating a clear description of the problem behavior (including frequency and intensity); identifying the antecedents, consequences, and other environmental factors that maintain the behavior; developing hypotheses that specify the motivating function of the behavior; and collecting direct observational data to test the hypothesis. Functional analysis also is useful in identifying antecedents and consequences that are associated with increased frequency of desirable behaviors so that they can be used to evoke new adaptive behaviors.

Structured Teaching
The TEACCH method, developed by Schopler and colleagues, emphasizes structure and has come to be called “structured teaching.” Important elements of structured teaching include organization of the physical environment, predictable sequence of activities, visual schedules, routines with flexibility, structured work/activity systems, and visually structured activities. There is an emphasis on both improving skills of individuals with ASDs and modifying the environment to accommodate their deficits. Several reports have documented progress in children who have received TEACCH services as well as parent satisfaction and improvement in parent teaching skills, but these reports were not from controlled studies of treatment outcomes. In a controlled trial, Ozonoff and Cathcart found that children treated with a TEACCH-based home program for 4 months in addition to their local day treatment programs improved significantly more than children in the control group who received local day treatment services only.

Developmental Models
Developmental models are based on use of developmental theory to organize hypotheses regarding the fundamental nature of ASDs and design approaches to address...
the deficits. The Denver model, for example, is based largely on remediating key deficits in imitation, emotion sharing, theory of mind, and social perception by using play, interpersonal relationships, and activities to foster symbolic thought and teach the power of communication. This program has shifted from a center-based treatment unit to service delivery in homes and inclusive school environments. Several studies have demonstrated improvements in cognitive, motor, play, and social skills beyond what would be expected on the basis of initial developmental rates in children who are treated according to the Denver model, but controlled trials are lacking.51–54

Relationship-focused early intervention models include Greenspan and Wieder’s developmental, individual-difference, relationship-based (DIR) model,55 Gutstein and Sheely’s relationship-development intervention (RDI),56 and the responsive-teaching (RT) curriculum developed by Mahoney et al.57,58 The DIR approach focuses on (1) “floor-time” play sessions and other strategies that are purported to enhance relationships and emotional and social interactions to facilitate emotional and cognitive growth and development and (2) therapies to remediate “biologically based processing capacities,” such as auditory processing and language, motor planning and sequencing, sensory modulation, and visual-spatial processing. Published evidence of the efficacy of the DIR model is limited to an unblinded review of case records (with significant methodologic flaws, including inadequate documentation of the intervention, comparison to a suboptimal control group, and lack of documentation of treatment integrity and how outcomes were assessed by informal procedures)55 and a descriptive follow-up study of a small subset (8%) of the original group of patients.59 RDI focuses on activities that elicit interactive behaviors with the goal of engaging the child in a social relationship so that he or she discovers the value of positive interpersonal activity and becomes more motivated to learn the skills necessary to sustain these relationships.56 Some reviewers have praised the face validity of this model, which targets the core impairment in social reciprocity. However, the evidence of efficacy of RDI is anecdotal; published empirical scientific research is lacking at this time. One study reported beneficial effects of RT on young children with ASDs or other developmental disabilities.58 Parents were taught to use RT strategies to encourage their children to acquire and use pivotal developmental behaviors (attention, persistence, interest, initiation, cooperation, joint attention, and affect). Children in both groups improved significantly on nonstandardized play-based measures of cognition and communication and standardized parent ratings of socioemotional functioning. Although a control group was lacking and the potential role of concurrent educational services was unclear, the improvements were beyond what the authors expected from maturational factors alone.58

Speech and Language Therapy
A variety of approaches have been reported to be effective in producing gains in communication skills in children with ASDs.9,17,20 Didactic and naturalistic behavioral methodologies (eg, DTT, verbal behavior, natural language paradigm, pivotal response training, milieu teaching) have been studied most thoroughly, but there is also some empirical support for developmental-pragmatic approaches (eg, Social Communication Emotional Regulation Transactional Support, Denver model, RDI, Hanen model).

People with ASDs have deficits in social communication, and treatment by a speech-language pathologist usually is appropriate. Most children with ASDs can develop useful speech, and chronologic age, lack of typical prerequisite skills, failure to benefit from previous language intervention, and lack of discrepancy between language and IQ scores should not exclude a child from receiving speech-language services.60 However, traditional, low-intensity pull-out service delivery models often are ineffective, and speech-language pathologists are likely to be most effective when they train and work in close collaboration with teachers, support personnel, families, and the child’s peers to promote functional communication in natural settings throughout the day.60

The use of augmentative and alternative communication modalities, including gestures, sign language, and picture communication programs, often is effective in enhancing communication.17,20,64 The Picture Exchange Communication System (PECS)62,63 is used widely. The PECS method incorporates ABA and developmental-pragmatic principles, and the child is taught to initiate a picture request and persist with the communication until the partner responds. Some nonverbal people with ASDs may benefit from the use of voice-output communication aids, but published evidence for these aids is scant.20,64 Introduction of augmentative and alternative communication systems to nonverbal children with ASDs does not keep them from learning to talk, and there is some evidence that they may be more stimulated to learn speech if they already understand something about symbolic communication.61,62,65

Social Skills Instruction
There is some objective evidence to support traditional and newer naturalistic behavioral strategies and other approaches to teaching social skills.22,24,66–68 Joint attention training may be especially beneficial in young, preverbal children with ASDs, because joint attention behaviors preceed and predict social language development.69,70 A recent randomized, controlled trial demonstrated that joint attention and symbolic play skills can be taught and that these skills generalize to different
settings and people. Families can facilitate joint attention and other reciprocal social interaction experiences throughout the day in the child’s regular activities. Examples of these techniques are described in the American Academy of Pediatrics parent booklet “Understanding Autism Spectrum Disorders.”

A social skills curriculum should target responding to the social overtures of other children and adults, initiating social behavior, minimizing stereotyped perseverative behavior while using a flexible and varied repertoire of responses, and self-managing new and established skills. Social skills groups, social stories, visual cuing, social games, video modeling, scripts, peer-mediated techniques, and play and leisure curricula are supported primarily by descriptive and anecdotal literature, but the quantity and quality of research is increasing. A number of social skills curricula and guidelines are available for use in school programs and at home.

**Occupational Therapy and Sensory Integration Therapy**

Traditional occupational therapy often is provided to promote development of self-care skills (eg, dressing, manipulating fasteners, using utensils, personal hygiene) and academic skills (eg, cutting with scissors, writing). Occupational therapists also may assist in promoting development of play skills, modifying classroom materials and routines to improve attention and organization, and providing prevocational training. However, research regarding the efficacy of occupational therapy in ASDs is lacking. Sensory integration (SI) therapy often is used alone or as part of a broader program of occupational therapy for children with ASDs. The goal of SI therapy is not to teach specific skills or behaviors but to remediate deficits in neurologic processing and integration of sensory information to allow the child to interact with the environment in a more adaptive fashion. Unusual sensory responses are common in children with ASDs, but there is not good evidence that these symptoms differentiate ASDs from other developmental disorders, and the efficacy of SI therapy has not been demonstrated objectively. Available studies are plagued by methodologic limitations, but proponents of SI note that higher-quality SI research is forthcoming. “Sensory” activities may be helpful as part of an overall program that uses desired sensory experiences to calm the child, reinforce a desired behavior, or help with transitions between activities.

**Comparative Efficacy of Educational Interventions for Young Children**

All treatments, including educational interventions, should be based on sound theoretical constructs, rigorous methodologies, and empirical studies of efficacy. Proponents of behavior analytic approaches have been the most active in using scientific methods to evaluate their work, and most studies of comprehensive treatment programs that meet minimal scientific standards involve treatment of preschoolers using behavioral approaches. However, there is still a need for additional research, including large controlled studies with randomization and assessment of treatment fidelity. Empirical scientific support for developmental models and other interventions is more limited, and well-controlled systematic studies of efficacy are needed.

Most educational programs available to young children with ASDs are based in their communities, and often, an “eclectic” treatment approach is used, which draws on a combination of methods including applied behavior analytic methods such as DTT; structured teaching procedures; speech-language therapy, with or without picture communication or related augmentative or alternative communication strategies; SI therapy; and typical preschool activities. Three studies that compared intensive ABA programs (25–40 hours/week) to equally intensive eclectic approaches have suggested that ABA programs were significantly more effective. Another study that involved children with ASDs and global developmental delay/mental retardation prospectively compared a less intensive ABA program (mean: 12 hours) to a comparably intensive eclectic approach and found statistically significant but clinically modest outcomes that favored those in the ABA group. Although the groups of children were similar on key dependent measures before treatment began, these studies were limited because of parent-determined rather than random assignment to treatment group. Additional studies to evaluate and compare educational treatment approaches are warranted.

**Programs for Older Children and Adolescents**

Some model programs provide programming throughout childhood and into adulthood. More commonly, the focus of specialized programs is on early childhood, and published research evaluating comprehensive educational programs for older children and adolescents with ASDs is lacking. However, there is empirical support for the use of certain educational strategies, particularly those that are based on ABA, across all age groups to increase and maintain desirable adaptive behaviors, reduce interfering maladaptive behaviors or narrow the conditions under which they occur, teach new skills, and generalize behaviors to new environments or situations.

When children with ASDs move beyond preschool and early elementary programs, educational intervention continues to involve assessment of existing skills, formulation of individualized goals and objectives, selection and implementation of appropriate intervention strategies and supports, assessment of progress, and adaptation of teaching strategies as necessary to enable students to acquire target skills. The focus on achieving social communication competence, emotional and be-
behavioral regulation, and functional adaptive skills necessary for independence continues. Educational programs should be individualized to address the specific impairments and needed supports while capitalizing on the child’s assets rather than being based on a particular diagnostic label.

Specific goals and objectives and the supports that are required to achieve them are listed in a child’s individualized education plan and should be the driving force behind decisions regarding the most appropriate, least restrictive classroom placement. Appropriate settings may range from self-contained special education classrooms to full inclusion in regular classrooms. Often, a mix of specialized and inclusive experience is appropriate. Even highly functioning students with ASDs often require accommodations and other supports such as provision of explicit directions, modification of classroom and homework assignments, organizational supports, access to a computer and word-processing software for writing tasks, and social communication skills training. When a paraprofessional aide is assigned, it is important that there be an infrastructure of expertise and support for the child beyond the immediate presence of the aide. The specific duties of the aide should be outlined, the strategies to be used should be defined, and the aide should receive adequate training.

In adolescence, the term “transition” is used to describe the movement from child-centered activities to adult-oriented activities. The major transitions are from the school environment to the workplace and from home to community living. In schools, transition-planning activities may begin at as early as 14 years of age, and by 16 years of age, the individualized education plan should include an individualized transition plan. The emphasis may shift from academic to vocational services and from remediating deficits to fostering abilities. A vocational assessment is often conducted to evaluate the adolescent’s interests and strengths and to determine the services and supports needed to promote independence in the workplace and in the community. Comprehensive transition planning involves the student, parents, teachers, the medical home, and representatives from all concerned community agencies. Depending on the individual’s cognitive level, social skills, health condition, work habits, and behavioral challenges, preparation for competitive, supported, or sheltered employment is targeted. Regardless of the type of employment, attention to skill development should never stop. Skills necessary for independent living should be taught to the degree possible given the abilities of the person. Sexuality education instruction should be included, and there is a growing body of literature that has addressed the topic.

**MEDICAL MANAGEMENT**

Children with ASDs have the same basic health care needs as children without disabilities and benefit from the same health-promotion and disease-prevention activities, including immunizations. In addition, they may have unique health care needs that relate to underlying etiologic conditions, such as fragile X syndrome or tuberous sclerosis, or to other conditions, such as epilepsy, that often are associated with ASDs. Those with pica or persistent mouthing of fingers or objects should be monitored for elevated blood lead concentrations, particularly if the history suggests potential for environmental exposure. These health care needs are most appropriately met within the context of a medical home.

To deliver appropriate and effective medical care, the history, approach to the patient, physical evaluation, and treatment options must be considered in the context of the patient’s ASD. Familiarizing the patient with the office setting and staff, allowing ample time while talking before touching the patient, allowing the child to manipulate instruments and materials, keeping instructions simple, using visual cues and supports, slowing down the pace, exaggerating social cues, and having family and/or familiar staff available may be helpful in reducing the obstacles to health care provision presented by patients’ difficulties with social interaction, communication, and accepting novelty. Often, more time is required for outpatient appointments. In a nationally representative sample, it was found that children with ASDs spent twice as much time with the physician per outpatient visit compared with children in control groups.

**Associated Morbidity and Mortality**

Health care utilization and costs are substantially higher for children and adolescents with ASDs compared with children without ASDs, and available data suggest that mortality is increased as well (standardized mortality ratio: 2.4–2.6). The increased mortality in ASDs is thought to be largely, but not completely, accounted for by the increased mortality associated with mental retardation and epilepsy. Cases of suicide in higher-functioning individuals have been reported.

**Seizures**

The reported prevalence of epilepsy among people with ASDs ranges from 11% to 39%. Comorbid severe global developmental delay/mental retardation and motor deficits are associated with a high prevalence of seizures (42%), whereas the prevalence of seizures is only 6% to 8% in children with ASDs without severe mental retardation, a motor deficit, an associated etiologic medical disorder, or a positive family history of epilepsy. The prevalence of epilepsy also was higher in studies that included adolescents and young adults, because the onset of epilepsy in ASDs has 2 peaks: 1 before 5 years of age and another in adolescence. Anticonvulsant treatment in children with ASDs is based on the same criteria that are used for other children with...
epilepsy, including accurate diagnosis of the particular seizure type.98

Epileptiform abnormalities on electroencephalography (EEG) are common in children with ASDs, with reported frequencies ranging from 10% to 72%.99 Some studies have suggested that epileptiform abnormalities on EEG100 and/or epilepsy101 are more common in the subgroup of children with ASDs who have a history of regression, whereas other studies have not demonstrated this association.102,103 Autistic regression with epileptiform abnormalities on EEG has been compared by analogy with Landau-Kleffner syndrome and electrical status epilepticus in sleep, but there are important differences between these conditions, and the treatment implications are unclear.94,104 Whether subclinical seizures have adverse effects on language, cognition, and behavior is debated, and there is no evidence-based recommendation for the treatment of children with ASDs and epileptiform abnormalities on EEG, with or without regression.104 Universal screening of patients with ASDs by EEG in the absence of a clinical indication is not currently supported.2,99 However, because of the increased prevalence of seizures in this population, a high index of clinical suspicion should be maintained, and EEG should be considered when there are clinical spells that might represent seizures.

Gastrointestinal Problems
The relationship between gastrointestinal problems and ASDs is unclear, because most studies have not examined representative groups of children with ASDs compared with appropriate controls.105,106 Surveys published in the gastroenterology literature have stated that gastrointestinal problems, such as chronic constipation or diarrhea, occur in 46% to 85% of children with ASDs.107–109 Lower rates in the range of 17% to 24% have been reported in other population-based studies,110–112 and a nested case-control study in the United Kingdom found that only 9% of children with ASDs and the same percentage of controls had a history of gastrointestinal complaints.113 However, in a recent cross-sectional study that used structured interviews and matched control groups, a lifetime history of gastrointestinal symptoms (including abnormal stool pattern, frequent constipation, frequent vomiting, and frequent abdominal pain) was elicited in 70% of the children with ASDs, compared with 42% of the children with other developmental disabilities (P = .03) and 28% of the children without developmental disabilities (P < .001).114

In children with ASDs undergoing endoscopy, who may or may not be representative of the general population of children with ASDs, high rates of lymphoid nodular hyperplasia and, often, histologically subtle esophagitis, gastritis, duodenitis, and colitis have been described, and preliminary evidence suggests that some immunohistochemical features may be unique to inflammation associated with ASDs.105,115,116 The existing literature does not support routine specialized gastroenterological testing for asymptomatic children with ASDs.105 However, if a child with an ASD presents with symptoms such as chronic or recurrent abdominal pain, vomiting, diarrhea, or constipation, it is reasonable to evaluate the gastrointestinal tract. Occult gastrointestinal discomfort also should be considered in a child who presents with a change in behavior, such as outbursts of aggression or self-injury. Radiographic evidence of constipation has been found to be more common in children with ASDs than in controls with abdominal pain (36% vs 10%),117 and effective management may provide global benefit.

Sleep Disturbance
Sleep problems are common in children and adolescents with ASDs at all levels of cognitive functioning.118–122 Sleep problems correlate with family distress and may have significant effects on daytime functioning and quality of life of children with ASDs.123–125 In some cases, there may be an identifiable etiology such as obstructive sleep apnea or gastroesophageal reflux; assessment and treatment are guided by history and physical examination. When there is not an identifiable medical cause, behavioral interventions including sleep-hygiene measures, restriction of daytime sleep, positive bedtime routines, and extinction procedures often are effective.118,126–129

Relatively little empirical information is available regarding pharmacologic management of sleep problems in children with ASDs or other developmental disabilities. Recommendations typically are based on case reports and open-label trials, extrapolation from the adult literature, and expert consensus (Table 1).128 There is some evidence of abnormality of melatonin regulation in children with ASDs125,130 and melatonin may be effective for improving sleep onset in children with ASDs as well as children with other developmental disabilities131–134 and otherwise healthy children with sleep/wake disorders.135 A recent open-label study suggested that controlled-release melatonin improved sleep in a group of 25 children with ASDs and that treatment gains were maintained at 1- and 2-year follow-up,136 but randomized, double-blind, placebo-controlled studies are needed. Recently, a child and a young adult with ASDs with significant insomnia were reported to have responded well, with no apparent adverse effects, to open-label treatment with the high-affinity melatonin receptor agonist ramelteon.137 Antihistamines, α2-agonists, benzodiazepines, chloral hydrate, trazodone, and newer nonbenzodiazepine hypnotic agents, such as zolpidem and zaleplon, sometimes are used to treat pediatric insomnia.128 In some cases, other conditions or symptoms, such as epilepsy, depression, anxiety, or aggressive outbursts, warrant pharmacologic treatment, and an agent that also may assist with sleep can be chosen.118
Evaluation of Challenging Behaviors

Problematic emotional reactions and behaviors such as aggression and self-injury are common in children and adolescents with ASDs. In some cases, medical factors may cause or exacerbate maladaptive behaviors, and recognition and treatment of medical conditions may eliminate the need for psychopharmacologic agents. For example, in the case of an acute onset or exacerbation of aggressive or self-injurious behavior, a source of pain or discomfort may be identified and treated.138 Sources of discomfort may include otitis media, otitis externa, pharyngitis, sinusitis, dental abscess, constipation, urinary

<table>
<thead>
<tr>
<th>Target Symptom Cluster</th>
<th>Potential Coexisting Diagnoses</th>
<th>Selected Medication Considerations</th>
<th>Selected References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repetitive behavior, behavioral rigidity, obsessive-compulsive symptoms</td>
<td>Obsessive-compulsive disorder, stereotypic movement disorder</td>
<td>SSRIs (fluoxetine, fluvoxamine, citalopram, escitalopram, paroxetine, sertraline)</td>
<td>McDougle et al158,b Buchsbaum et al159,b Sugie et al160,b Hollander et al161,b Moore et al162,b Namerow et al163,b Owley et al164,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Atypical antipsychotic agents (risperidone, aripiprazole, olanzapine, quetiapine, ziprasidone)</td>
<td>McDougle et al165,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Valproic acid†</td>
<td></td>
</tr>
<tr>
<td>Hyperactivity, impulsivity, inattention</td>
<td>Attention-deficit/hyperactivity disorder</td>
<td>Stimulants (methylphenidate, dextroamphetamine, mixed amphetamine salts)</td>
<td>Quintana et al166,b Handen et al167,b RUPP Autism Network168,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α2-agonists (clonidine, guanfacine)</td>
<td></td>
</tr>
<tr>
<td>Aggression, explosive outbursts, self-injury</td>
<td>Intermittent explosive disorder</td>
<td>Atypical antipsychotic agents (risperidone, aripiprazole, olanzapine, quetiapine, ziprasidone)</td>
<td>McDougle et al161,b Arnold et al162,b Shea et al163,b RUPP Autism Network164,b Troost et al165,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α2-agonists (clonidine, guanfacine)</td>
<td></td>
</tr>
<tr>
<td>Sleep dysfunction</td>
<td>Circadian rhythm sleep disorder, dyssomnia—not otherwise specified</td>
<td>Anticonvulsant mood stabilizers (levetiracetam, topiramate, valproic acid)</td>
<td>Hollander et al168,b Rugino and Samsock169,b Hardan et al170,b, Myers171,b, Myers and Chalmann172,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSRIs (fluoxetine, fluvoxamine, citalopram, escitalopram, paroxetine, sertraline)</td>
<td>McDougle et al173,b Moore et al174,b Namerow et al175,b Owley et al176,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β-blockers (propranolol, nadolol, metoprolol, pindolol)</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>Generalized anxiety disorder, anxiety disorder—not otherwise specified</td>
<td>Ramelteon</td>
<td>Connor et al177,b Ratey et al178,b Myers and Chalmann179,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antihistamines (diphenhydramine, hydroxyzine)</td>
<td>Giannotti et al180,b Jan and Freeman181,b Phillips and Appleton182,b Turk183,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α2-agonists (clonidine, guanfacine)</td>
<td>Owens et al184,b Stigler et al185,b Reed and Findling186,b Owens et al187,b</td>
</tr>
<tr>
<td>Depressive phenotype (marked change from baseline including symptoms such as social withdrawal, irritability, sadness or crying spells, decreased energy, anorexia, weight loss, sleep dysfunction)</td>
<td>Major depressive disorder, depressive disorder—not otherwise specified</td>
<td>SSRIs (fluoxetine, fluvoxamine, citalopram, escitalopram, paroxetine, sertraline)</td>
<td>McDougle et al188,b Buchsbaum et al189,b Sugie et al190,b Hollander et al191,b Moore et al192,b Namerow et al193,b Owley et al194,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Buspirone</td>
<td>Butelaar et al195,b Posey et al196,b</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mirtazapine</td>
<td>McDougle et al197,b Moore et al198,b Namerow et al199,b Owley et al200,b</td>
</tr>
</tbody>
</table>

---

**TABLE 1 Selected Potential Medication Options for Common Target Symptoms or Coexisting Diagnoses in Children With ASDs**

*Note: *The table above lists potential medication options for common target symptoms or coexisting diagnoses in children with ASDs. The selected medications are based on evidence from various studies and clinical trials. Each option includes the targeted condition, the medication classes considered, and the references for further information. The table is not exhaustive and additional research is recommended for comprehensive treatment decisions.*
Extreme food selectivity has the potential to lead to protein-calorie malnutrition or specific vitamin or mineral deficiencies; however, most studies that evaluated nutritional status in children with ASDs have suggested that despite dietary selectivity, malnutrition is uncommon. Although the prevalence in children with ASDs is unknown, pica related to iron or zinc deficiency may respond to supplementation with the appropriate mineral. It should be noted that it is not clear how frequently medical factors cause or exacerbate serious maladaptive behaviors in children with ASDs, and the efficacy of these interventions is based on anecdotes, case reports, and conventional clinical practice rather than empirical support from clinical trials.

It is also important to consider environmental factors that may precipitate challenging behaviors. Parents, teachers, or other caregivers may inadvertently reinforce maladaptive behaviors, and in such cases, the most appropriate and effective interventions are behavioral. In some instances, a mismatch between educational or behavioral expectations and cognitive ability of the child is responsible for disruptive behavior (eg, when the diagnosis of mental retardation has not been recognized), and adjustment of expectations is the most appropriate intervention. In both situations, a functional analysis of behavior, completed by a behavior specialist in the settings in which the problems occur, will identify factors in the environment that exacerbate or maintain the problematic behavior. A strategy for intervention through behavioral techniques and environmental manipulations can then be formulated and tested.

**Psychopharmacology**

Pharmacologic interventions may be considered for maladaptive behaviors such as aggression, self-injurious behavior, repetitive behaviors (eg, perseveration, obsessions, compulsions, and stereotypic movements), sleep disturbance, mood lability, irritability, anxiety, hyperactivity, inattention, destructive behavior, or other disruptive behaviors. After treatable medical causes and modifiable environmental factors have been ruled out, a therapeutic trial of medication may be considered if the behavioral symptoms cause significant impairment in functioning and are suboptimally responsive to behavioral interventions. In some cases, the diagnosis of a comorbid disorder, such as major depression, bipolar disorder, or an anxiety disorder, can be made reasonably and the patient can be treated with medications that are useful for treating these conditions in otherwise typically developing children and adolescents. Modifications of diagnostic criteria may be necessary to account for clinical presentations of psychiatric conditions in individuals with developmental disabilities, and tools such as behavior checklists and structured interviews may be helpful. In other cases, clinicians opt to target specific interfering maladaptive behaviors or symptom clusters in the absence of a clear comorbid psychiatric diagnosis (a target-symptom approach).

Recent surveys indicate that approximately 45% of children and adolescents and up to 75% of adults with ASDs are treated with psychotropic medication. Greater age, lower adaptive skills and social competence, and higher levels of challenging behavior are associated with the likelihood of medication use. The evidence regarding the efficacy of psychopharmacologic interventions in patients with ASDs has been de-
tained in recent reviews. Although most psychotropic medications have been used in children with ASDs, there is currently insufficient literature to establish consensus regarding an evidence-based approach to pharmacologic management. However, in recent years, there has been an increase in publication of randomized, double-blind, placebo-controlled clinical trials to guide clinical practice.

Surveys performed in the United States suggest that selective serotonin-reuptake inhibitors (SSRIs), atypical antipsychotic agents, stimulants, and α2-adrenergic agonist antihypertensive agents are the most commonly prescribed classes of psychotropic medications for children with ASDs. Double-blind, placebo-controlled trials have demonstrated efficacy of the SSRIs fluoxetine and fluvoxamine in the treatment of repetitive and other maladaptive behaviors in patients with ASDs, and open-label trials of these and other SSRIs have shown improvements in target symptoms, including repetitive behaviors, irritability, depressive symptoms, tantrums, anxiety, aggression, difficulty with transitions, and aspects of social interaction and language. Potential adverse effects of SSRIs include but are not limited to nausea, drowsiness, sexual dysfunction, constipation, abdominal discomfort, fatigue, dizziness, dry mouth, agitation, behavioral activation, hypomania or mania, apathy, suicidal ideation, and alteration of sleep.

Risperidone has become the first medication with US Food and Drug Administration–approved labeling for the symptomatic treatment of irritability (including aggressive behavior, deliberate self-injury, and temper tantrums) in children and adolescents with ASDs. Two large, multisite, randomized, controlled trials have confirmed the short-term efficacy of risperidone for these severe disruptive behaviors in youth with ASDs, and 2 open-label studies, each with a double-blind discontinuation component, have suggested long-term benefits and tolerance. Potential adverse effects include but are not limited to appetite reduction, inhibition of growth, delayed sleep onset, jitteriness, exacerbation of tics, abdominal discomfort, increased blood pressure, increased heart rate, irritability, increased anxiety, and repetitive behaviors.

Two small double-blind, placebo-controlled trials have documented modest benefits of clonidine in reducing hyperarousal symptoms including hyperactivity, irritability and outbursts, impulsivity, and repetitive behaviors in children with ASDs. A prospective open-label trial and a retrospective record review have suggested that guanfacine was similarly effective in some patients. Potential adverse effects of these centrally acting α2-agonists include but are not limited to dry mouth, agitation, behavioral activation, dizziness, constipation, and irritability, and these drugs can be dangerous in overdose. Recently, a retrospective study, an open-label trial, and a small double-blind, placebo-controlled crossover trial suggested that atomoxetine may be effective for attention-deficit/hyperactivity disorder–like symptoms in children and adolescents with ASDs, and additional research is warranted. Appetite suppression, nausea, fatigue, mood swings, suicidal ideation, dizziness, and liver injury are among the potential adverse effects of atomoxetine.

A summary of selected target symptoms, potential psychiatric diagnoses, and medication options is provided in Table 1. Pediatricians and other practitioners should only prescribe medications with which they have sufficient expertise, including knowledge of indications and contraindications, dosing, potential adverse effects, drug-drug interactions, and monitoring requirements. It will be important for future research to address the need for more rigorous evaluation of safety and efficacy of psychotropic agents in children with ASDs; the value of combining behavioral and medical interventions; the practice of polypharmacy; delineation of clinical and biological subgroups of patients who may be responsive to particular treatments; the role of drugs in treating deficits in language and nonlanguage cognition, social interaction, and behavioral rigidity; and the potential to alter the neural substrate during early critical periods to affect brain development and future function. Several multisite trials are underway, and others undoubtedly will be forthcoming.

Principles to guide the approach to psychopharmacologic management of ASDs in clinical practice have been proposed by several authors in recent years, and an approach is outlined in Table 2. When medications are used, potential benefits and adverse effects should be explained, informed consent should be obtained, baseline data regarding behaviors and somatic complaints should be collected, and potential strategies for dealing with treatment failure or partial response should be reviewed. It is important to have some quantifiable means of assessing the efficacy of the medication and to
TABLE 2  Clinical Approach to Psychopharmacologic Management

| Identify and assess target behaviors |
| Parent/caregiver interview |
| Intensity |
| Duration |
| Exacerbating factors/triggers (time, setting/location, demand situations, denials, transitions, etc) |
| Ameliorating factors and response to behavioral interventions |
| Time trends (increasing, decreasing, stable) |
| Degree of interference with functioning |
| Consider baseline behavior-rating scales and/or baseline performance measures/direct observational data |
| Include input from school staff and other caregivers |
| Assess existing and available supports |
| Behavioral services and supports |
| Educational program, habilitative therapies |
| Respite care, family psychosocial supports |
| Search for medical factors that may be causing or exacerbating target behavior(s) |
| Consider sources of pain or discomfort (infectious, gastrointestinal, dental, allergic, etc) |
| Consider other medical causes or contributors (sleep disorders, seizures, menstrual cycle, etc) |
| Complete any medical tests that may have a bearing on treatment choice |
| Consider psychotropic medication on the basis of the presence of |
| Evidence that the target symptoms are interfering substantially with learning/academic progress, socialization, health/safety (of the patient and/or others around him or her), or quality of life |
| Suboptimal response to available behavioral interventions and environmental modifications |
| Research evidence that the target behavioral symptoms or coexisting psychiatric diagnoses are amenable to pharmacologic intervention |
| Choose a medication on the basis of |
| Likely efficacy for the specific target symptoms |
| Potential adverse effects |
| Practical considerations such as formulations available, dosing schedule, cost, and requirement for laboratory or electrocardiographic monitoring |
| Informed consent (verbal or written) from parent/guardian and, when possible, assent from the patient |
| Establish plan for monitoring of effects |
| Identify outcome measures |
| Discuss time course of expected effects |
| Arrange follow-up telephone contact, completion of rating scales, reassessment of behavioral data, and visits accordingly |
| Outline a plan regarding what might be tried next if there is a negative or suboptimal response or to address additional target symptoms |
| Change to a different medication |
| Add another medication to augment a partial or suboptimal therapeutic response to the initial medication (same target symptoms) |
| Add a different medication to address additional target symptoms that remain problematic |
| Obtain baseline laboratory data if necessary for the drug being prescribed and plan appropriate follow-up monitoring |
| Explore the reasonable dose range for a single medication for an adequate length of time before changing to or adding a different medication |
| Monitor for adverse effects systematically |
| Consider careful withdrawal of the medication after 6–12 mo of therapy to determine whether it is still needed |


obtain input from a variety of sources, such as parents, teachers, therapists, and aides. Consistent use of validated, treatment-sensitive rating scales and medication adverse-effect scales is desirable. A wide variety of outcome measures have been used in research trials and in clinical practice to measure maladaptive behavior treatment effects. Among the most common are the Clinical Global Impression Scale, Aberrant Behavior Checklist, and Nisonger Child Behavior Rating Form.

Complementary and Alternative Medicine

Complementary and alternative medicine (CAM) is defined by the National Center for Complementary and Alternative Medicine as “a group of diverse medical and health care systems, practices, and products that are not presently considered to be part of conventional medicine.” The definition of CAM adopted by the Cochrane Collaboration is “a broad domain of healing resources that encompasses all health systems, modalities, and practices and their accompanying theories and beliefs, other than those intrinsic to the politically dominant health systems of a particular society or culture in a given historical period.” Detailed reviews of CAM as related to developmental disabilities and ASD-specific CAM have been published recently.

Use of CAM is common in children with ASDs. Levy et al reported that by the time their clinical population received a formal diagnostic evaluation for a suspected ASD, almost one third of the children already had received a complementary or alternative therapy, and a survey conducted in a predominantly white, middle-to-upper socioeconomic-status private-practice population found that 92% of parents who responded had used CAM therapies for their children with ASDs.
Another recent parent survey found that 52% of the children with an ASD had been treated with at least 1 CAM therapy, compared with 28% of a group of control children without disabilities.207 Surveys indicate that only 36% to 62% of caregivers who used CAM therapies for their children with ASDs had informed the child’s primary care physician,207,208 although more information on CAM is something that families indicate that they want from their child’s primary health care professionals.209

It is important that health care professionals understand how to evaluate the evidence used to support all treatments, including CAM, psychopharmacologic, and other interventions. Ideally, the evidence supporting or refuting a treatment should include peer-reviewed studies with appropriately diagnosed, well-defined homogeneous study populations; a randomized, double-blind, placebo-controlled design; an adequate sample size to support the statistical analysis presented; control for confounding factors; and use of appropriate, validated outcome measures. When evaluating the efficacy of studies, it is particularly important to keep in mind confounding factors, such as the placebo effect, and the natural history of the disorder. Participation in a study may alter the way a parent interacts with a child and confound the perceived outcome,210 and improvements are expected with maturation and educational interventions. Only appropriately controlled studies are helpful in proving that an effect is attributable to the intervention being studied.

The practitioner should encourage families to seek additional information when they encounter the following claims or situations:

- treatments that are based on overly simplified scientific theories;
- therapies that are claimed to be effective for multiple different, unrelated conditions or symptoms;
- claims that children will respond dramatically and some will be cured;
- use of case reports or anecdotal data rather than carefully designed studies to support claims for treatment;
- lack of peer-reviewed references or denial of the need for controlled studies; or
- treatments that are said to have no potential or reported adverse effects.

To help to describe their proposed rationales and mechanisms, CAM therapies used to treat ASDs have been categorized as “nonbiological” or “biological.”204 Examples of nonbiological interventions include treatments such as auditory integration training, behavioral optometry, craniosacral manipulation, dolphin-assisted therapy, music therapy, and facilitated communication. Examples of biological therapies include immunoregulatory interventions (eg, dietary restriction of food allergens or administration of immunoglobulin or antiviral agents), detoxification therapies (eg, chelation), gastrointestinal treatments (eg, digestive enzymes, antifungal agents, probiotics, “yeast-free diet,” gluten/casein-free diet, and vancomycin), and dietary supplement regimens that are purported to act by modulating neurotransmission or through immune factors or epigenetic mechanisms (eg, vitamin A, vitamin C, vitamin B6, and magnesium, folic acid, folinic acid, vitamin B12, dimethylglycine and trimethylglycine, carnosine, omega-3 fatty acids, inositol, various minerals, and others).203,204

For most of the aforementioned CAM interventions, there is not enough scientific evidence yet to support or refute their use as treatment for ASDs. However, evaluation of treatments is possible, and a few CAM treatments have been appropriately studied. For example, more than a dozen randomized, double-blind, placebo-controlled trials involving more than 700 patients have demonstrated that secretin (a biological treatment) is not an effective treatment for ASDs.212,213 Evaluation of nonbiological treatments also is feasible. This has been demonstrated in the case of facilitated communication, a technique that uses a trained facilitator to provide physical support to a nonverbal person’s hand or arm while that person uses a computer keyboard or other device to spell. Evidence suggests that the communications produced actually originate from the facilitator and that facilitated communication is not a valid treatment for ASDs.214–218

Because of methodologic flaws, insufficient numbers of patients, or lack of replication, many CAM therapies have been inadequately evaluated; therefore, evidence-based recommendations for their use are not possible. The most recent and most appropriately designed trials have demonstrated no significant benefit of dimethylglycine,219,220 vitamin B6, and magnesium,221,222 or auditory integration training.223–225 Both positive and negative results have been described for small, methodologically flawed studies of intravenous immunoglobulin. A recent double-blind, placebo-controlled trial revealed no statistically significant differences on Aberrant Behavior Checklist subscale scores between small groups of children with ASDs who were given omega-3 fatty acids and those who were given placebo.229 However, the investigators noted a trend toward superiority of omega-3 fatty acids over placebo for hyperactivity, which suggests that further investigation may be warranted.229 The gluten/casein-free diet is based on a hypothesis of abnormal gut permeability and exogenous opiate excess. Although use of the gluten/casein-free diet for children with ASDs is popular, there is little evidence to support or refute this intervention, and reviewers have determined that meaningful conclusions cannot be drawn from the existing literature.230,231 Subsequent to these reviews, a randomized, double-blind
pilot study demonstrated no significant benefit. Double-blind, placebo-controlled elimination and challenge studies are in progress, and it is anticipated that these studies will provide substantially more useful information regarding the efficacy of the gluten/casein-free diet. Measurement of urinary peptides has not been proven to be clinically useful as a marker for ASDs or as a tool to determine if dietary restriction is warranted or would be effective.

Many popular interventions, such as chelation of heavy metals, antifungal agents to decrease presumed yeast overgrowth, and antiviral agents to modulate the immune system, have not yet been studied in people with ASDs; their popularity is based on unproven theories and anecdotes or case reports. None of these interventions can be endorsed as treatment for ASDs outside of well-designed and appropriately monitored clinical trials. Some treatments, such as intravenous chelation, may be particularly dangerous and should be discouraged. One child with autism died as a result of chelation with edetate disodium (Na2EDTA) despite the facts that a causal association between mercury and ASDs has not been demonstrated, there is no scientific evidence that chelation is an effective treatment for ASDs, and the effectiveness of chelation therapy to improve nervous system symptoms of chronic mercury toxicity has not been established. Unless there is clear evidence of current heavy metal toxicity, chelation by any method is not indicated outside of monitored clinical trials.

In some cases, interesting findings await replication or further investigation. For example, in a double-blind, placebo-controlled trial of vitamin C, improvement was found in total and sensory motor scores on the Ritvo-Freeman Real Life Rating Scale and several small studies have suggested that music therapy had some short-term benefit on communication skills but not on behavior problems of children with ASDs. Recently, a group of 20 children with ASDs were compared with children without ASDs and found to have an imbalance of methionine and homocysteine metabolism, which was interpreted to represent impaired capacity for methylation and increased oxidative stress. Treatment with trimethylglycine, folic acid, and methylcobalamin resulted in normalization of laboratory findings. The study did not measure clinical response to the intervention, but anecdotal improvements were noted. Interpretation of these preliminary findings awaits further investigation.

Health care practitioners who diagnose and treat children with ASDs should recognize that many of their patients will use nonstandard therapies. The importance of becoming knowledgeable about CAM therapies, inquiring about current and past CAM use, providing balanced information and advice about treatment options, identifying risks or potential harmful effects, avoiding becoming defensive or dismissing CAM in ways that convey a lack of sensitivity or concern, maintaining open communication, and continuing to work with families even if there is disagreement about treatment choices has been emphasized. It also is essential to critically evaluate the scientific merits of specific therapies and share this information with families, educate families about how to evaluate information and recognize pseudoscience, and insist that studies that examine CAM be held to the same scientific and ethical standards as all clinical research.

Parents of children with ASDs will understandably pursue interventions that they believe may present some hope of helping their child, particularly if the therapies are viewed as being unlikely to have any adverse effects. Unfortunately, families are often exposed to unsubstantiated, pseudoscientific theories and related clinical practices that are, at best, ineffective and, at worst, compete with validated treatments or lead to physical, emotional, or financial harm. Time, effort, and financial resources expended on ineffective therapies can create an additional burden on families. Health care professionals can help parents and other caregivers to distinguish empirically validated treatment approaches from treatments that have been proven to be ineffective and those that are unproven and potentially ineffective and/or harmful.

FAMILY SUPPORT
Management should focus not only on the child but also on the family. Although parents once were viewed erroneously as the cause of a child’s ASD, it is now recognized that parents play a key role in effective treatment. Having a child with an ASD has a substantial effect on a family. Parents and siblings of children with ASDs experience more stress and depression than those of children who are typically developing or even those who have other disabilities. Supporting the family and ensuring its emotional and physical health is an extremely important aspect of overall management of ASDs.

Physicians and other health care professionals can provide support to parents by educating them about ASDs; providing anticipatory guidance; training and involving them as cotherapists; assisting them in obtaining access to resources; providing emotional support through traditional strategies such as empathetic listening and talking through problems; and assisting them in advocating for their child’s or sibling’s needs. In some cases, referral of parents for counseling or other appropriate mental health services may be required. The need for support is longitudinal, although the specific needs may vary throughout the family life cycle.

One of the chief strategies for helping families raise children with ASDs is helping to provide them with access to needed ongoing supports and additional services during critical periods and/or crises. Natural supports include spouses, extended family members, neigh-
bors, religious institutions, and friends who can help with caregiving and who can provide psychological and emotional support. Informal supports include social networks of other families of children with ASDs and community agencies that provide training, respite, social events, and recreational activities. Formal supports include publicly funded, state-administered programs such as early intervention, special education, vocational and residential/living services, respite services, Medicaid, in-home and community-based waiver services, Supplemental Security Income benefits, and other financial subsidies. The breadth and depth of services vary, even within the same state or region. Few services exist in many rural areas, and public programs may have long waiting lists.

Sibling support groups offer the opportunity to learn important information and skills while sharing experiences and connecting with other siblings of children with ASDs.244 Although the research on support groups for siblings of children with disabilities is difficult to interpret because of study-design problems and inconsistent outcome effects on sibling adjustment, these groups generally have been evaluated positively by participating siblings and parents,244 and there is some evidence of beneficial effects for siblings of children with ASDs.245

Because each state has organized its services and access mechanisms differently, physicians and families must learn their own state’s unique rules to access supports by contacting the state or county offices of the states’ Department of Health and Human Services or Mental Health and Mental Retardation or the state developmental disabilities organization. In addition, local parent advocacy organizations, national autism and related developmental disability organizations, early intervention administrators, and school district special education coordinators often are knowledgeable about various programs and their respective eligibility requirements.

CONCLUSIONS

ASDs are chronic conditions that affect nearly 1 of every 150 children and require ongoing medical and nonmedical intervention. There is a growing body of evidence that supports the efficacy of certain interventions in ameliorating symptoms and enhancing functioning, but much remains to be learned. In addition to their important roles in identifying ASDs through screening and surveillance, establishing the diagnosis, conducting an etiologic evaluation, and providing genetic counseling after a diagnosis is made,2 pediatricians and other primary health care professionals are in a position to provide important longitudinal medical care and to support and educate families and guide them to empirically supported interventions for their children.

COUNCIL ON CHILDREN WITH DISABILITIES EXECUTIVE COMMITTEE, 2006–2007
Paul H. Lipkin, MD, Chairperson
J. Daniel Cartwright, MD
Larry W. Desch, MD
John C. Duby, MD
Ellen Roy Elias, MD
Chris Plauché Johnson, MD, MEd
Eric B. Levey, MD
Gregory S. Liptak, MD
Nancy A. Murphy, MD
Scott M. Myers, MD
Ann Henderson Tilton, MD

LIAISONS
Donald Lollar, EdD
Centers for Disease Control and Prevention
Michelle Macias, MD
Section on Developmental and Behavioral Pediatrics
Merle McPherson, MD, MPH
Maternal and Child Health Bureau
Donna Gore Olson
Family Voices
Bonnie Strickland, PhD
Maternal and Child Health Bureau

STAFF
Stephanie Mucha Skipper, MPH
Jill Ackermann, MS
Mark Del Monte, JD

CONTRIBUTORS
Thomas D. Challman, MD
Susan L. Hyman, MD
Susan E. Levy, MD
S. Andrew Spooner, MD
Partnership for Policy Implementation
Marshalyn Yeargin-Allsopp, MD

REFERENCES
3. Fombonne E, Zakarian R, Bennett A, Meng L, McLean-Heywood D. Pervasive developmental disorders in Montreal, Quebec, Canada: prevalence and links with immunizations. Pediatrics. 2006;118(1). Available at: www.pediatrics.org/cgi/content/full/118/1/e139


5. Handlerman JS, Harris SL. Preschool Education Programs for Children With Autism. 2nd ed. Austin, TX: Pro-Ed; 2000


126. Christodulu KV, Durand VM. Reducing bedtime disturbance and night waking using positive bedtime routines and sleep restriction. Focus Autism Other Dev Disabl. 2004;19:130–139
156. Posey DJ, McDougle CJ. The pharmacotherapy of target symptoms associated with autistic disorder and other pervasive developmental disorders. Harv Rev Psychiatry. 2000;8:45–63


**RESOURCE FOR FAMILIES**

Identification and Evaluation of Children With Autism Spectrum Disorders

Chris Plauché Johnson, MD, MEd, Scott M. Myers, MD, and the Council on Children With Disabilities

ABSTRACT

Autism spectrum disorders are not rare; many primary care pediatricians care for several children with autism spectrum disorders. Pediatricians play an important role in early recognition of autism spectrum disorders, because they usually are the first point of contact for parents. Parents are now much more aware of the early signs of autism spectrum disorders because of frequent coverage in the media; if their child demonstrates any of the published signs, they will most likely raise their concerns to their child’s pediatrician. It is important that pediatricians be able to recognize the signs and symptoms of autism spectrum disorders and have a strategy for assessing them systematically. Pediatricians also must be aware of local resources that can assist in making a definitive diagnosis of, and in managing, autism spectrum disorders. The pediatrician must be familiar with developmental, educational, and community resources as well as medical subspecialty clinics. This clinical report is 1 of 2 documents that replace the original American Academy of Pediatrics policy statement and technical report published in 2001. This report addresses background information, including definition, history, epidemiology, diagnostic criteria, early signs, neuropathologic aspects, and etiologic possibilities in autism spectrum disorders. In addition, this report provides an algorithm to help the pediatrician develop a strategy for early identification of children with autism spectrum disorders. The accompanying clinical report addresses the management of children with autism spectrum disorders and follows this report on page 1162 [available at www.pediatrics.org/cgi/content/full/120/5/1162]. Both clinical reports are complemented by the toolkit titled “Autism: Caring for Children With Autism Spectrum Disorders: A Resource Toolkit for Clinicians,” which contains screening and surveillance tools, practical forms, tables, and parent handouts to assist the pediatrician in the identification, evaluation, and management of autism spectrum disorders in children.

INTRODUCTION

Public and physician awareness of autism has increased markedly in the new millennium because of increased media coverage and a rapidly expanding body of knowledge published in professional journals. Professionals who specialize in autism have proliferated over the past 2 decades and have introduced the terminology “autism spectrum disorders” (ASDs) to reflect the broader spectrum of clinical characteristics that now define autism.1,2 ASDs represent 3 of the pervasive developmental disorders defined in the Diagnostic and Statistical Manual of Mental Disorders, 4th edition, text revision (DSM-IV-TR)3 as autistic disorder (AD), Asperger syndrome (AS), and pervasive developmental disorder—not otherwise specified (PDD-NOS).4 What is termed autism spectrum disorders (ASDs) now includes all these diagnoses along with other clinical features that are necessary for the diagnosis of autism. This terminology reflects the continuous spectrum of clinical characteristics observed in individuals with autism.5

Key Words

autism, autism spectrum disorders, Asperger syndrome, pervasive developmental disorders, fragile X syndrome, joint attention, self-injurious behaviors, theory of mind, neuropathologic abnormalities

Abbreviations

ASD—autism spectrum disorder
AD—autistic disorder
DSM—Diagnostic and Statistical Manual of Mental Disorders
AS—Asperger syndrome
PDD-NOS—pervasive developmental disorder—not otherwise specified
PCP—primary care pediatrician
AAP—American Academy of Pediatrics
IDEA—Individuals With Disabilities Education Act
MR—mental retardation
GDD—global developmental delay
ADHD—attention-deficit/hyperactivity disorder
FISH—fluorescence in situ hybridization
MMR—measles-mumps-rubella
JA—joint attention
ToM—theory of mind
SLP—speech-language pathologist
CHAT—Checklist for Autism in Toddlers
M-CHAT, Modified Checklist for Autism in Toddlers
CAST—Childhood Asperger Syndrome Test
EEG—electroencephalography
Disorders, Fourth Edition (DSM-IV), and the newer Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR): autistic disorder (AD), Asperger syndrome (AS [this terminology will be used in this report, although “Asperger’s disorder” is used in the aforementioned publications]), and pervasive developmental disorder—not otherwise specified (PDD-NOS). In addition to being a spectrum disorder, autism has wide variability with respect to the presence and intensity of symptoms, even within the DSM-IV-TR categories, which indicates that there may be additional subtypes.

ASDs are not rare; many primary care pediatricians (PCPs) care for several children with ASDs. In fact, a survey completed in 2004 revealed that 44% of PCPs reported that they care for at least 10 children with ASDs; however, only 8% stated that they routinely screened for ASDs. Another survey indicated that although PCPs were aware of the current DSM-IV-TR diagnostic criteria, they sometimes held beliefs about ASDs that were outdated. It is critical that PCPs recognize the early signs of ASDs and be aware of new data that support better outcomes in children whose conditions are diagnosed early and who participate in appropriate intervention programs.

Because it is a chronic condition, the PCP also needs to feel comfortable with the ongoing care of children with ASDs within the context of the medical home. To support PCPs in the identification and care of children with ASD, the American Academy of Pediatrics (AAP) has developed and distributed several documents:

- The “Autism A.L.A.R.M.”: a flyer that highlights the prevalence of autism, the importance of screening and listening to parents’ concerns, and the urgency of making simultaneous referrals to specialists in ASDs and early intervention programs to promote improved outcomes.
- “Is Your One-Year-Old Communicating With You?”: a brochure that focuses on early identification of social communication deficits and behavior problems that may be associated with developmental disorders, primarily ASDs. This brochure is intended for distribution to all parents of infants at the 9- or 12-month well-child visit. It encourages parents to share any concerns they have about their infant’s language development and social skills with the pediatrician as early as possible.
- “Understanding Autism Spectrum Disorders”: a 48-page introductory booklet for parents of children in whom an ASD has been diagnosed recently or is suspected strongly.

In addition, the AAP has developed an ASD toolkit and resource guide to assist the PCP with implementation of the principles discussed herein.

Although ASDs are neurodevelopmental conditions with strong genetic underpinnings, their exact etiology is unknown. In 1943, Leo Kanner, a psychiatrist at Johns Hopkins University, first described autism in a small group of children who demonstrated extreme aloofness and total indifference to other people. In 1944, Hans Asperger, an Austrian pediatrician who was unaware of Kanner’s work, published an article that described children who demonstrated symptoms similar to those of Kanner’s patients, with the exception that verbal and cognitive skills were higher. The term “infantile autism” first appeared as a diagnostic label in the Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III). Since then, terminology has changed and diagnostic criteria have broadened. Diagnostic criteria for AS were not included in the DSM until the fourth edition (DSM-IV). The most recent criteria for AD and AS (Asperger’s disorder) are found in the DSM-IV-TR (Tables 1 and 2, respectively). PDD-NOS, the remaining ASD, is described in the DSM-IV-TR as a subthreshold diagnostic term used when a child demonstrates severe and pervasive impairments in reciprocal social skills associated with deficits in language skills or with the presence of stereotypic behaviors or restricted interests or activities but does not meet full criteria for AD or AS. Although Rett syndrome and childhood disintegrative disorder are included in the DSM-IV-TR listings, they are not considered ASDs but should be considered in the differential diagnosis of each child, depending on the presenting signs and symptoms.

**Epidemiology**

Authors of studies published early in the new millennium concluded that the best estimate of current prevalence of ASDs in Europe and North America is approximately 6 per 1000. In 2000, the Centers for Disease Control and Prevention organized the Autism and Developmental Disabilities Monitoring Network, a multisite, records-based surveillance program, to study the prevalence of ASDs. The network uses systematic screening of developmental evaluation records for autistic behaviors rather than depending on a medical or educational diagnostic label of an ASD. In 2007, the network reported ASD rates for 8-year-old children ranging from 1 in 303 to 1 in 94 for 2 time periods (2000 and 2002) in a total of 14 sites in the United States; the average rate was 1 in 150 or 6.6 per 1000 8-year-olds. Although these studies reflect a 10-fold increase from studies published a half-century ago that chiefly targeted AD alone, most of the newer studies also included individuals with AS and PDD-NOS. One of the few studies that analyzed the prevalence in regard to type of ASD revealed that in Canada, where the overall rate was 6.5 per 1000, the individual rates were 2.2 per 1000 for AD, 1.0 per 1000 for AS, and 3.3 per 1000 for PDD-NOS. Studies have varied in design, and...
TABLE 1  Diagnostic Criteria for 299.00: AD

A. A total of six (or more) items from (1), (2), and (3), with at least two from (1), and one each from (2) and (3):

(1) qualitative impairment in social interaction, as manifested by at least two of the following:
   (a) marked impairment in the use of multiple nonverbal behaviors such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction
   (b) failure to develop peer relationships appropriate to developmental level
   (c) a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (eg, by a lack of showing, bringing, or pointing out objects of interest)
   (d) lack of social or emotional reciprocity

(2) qualitative impairments in communication as manifested by at least one of the following:
   (a) delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime)
   (b) in individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others
   (c) stereotyped and repetitive use of language or idiosyncratic language
   (d) lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level

(3) restricted repetitive and stereotyped patterns of behavior, interests, and activities, as manifested by at least one of the following:
   (a) encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus
   (b) apparently inflexible adherence to specific, nonfunctional routines or rituals
   (c) stereotyped and repetitive motor mannerisms (eg, hand or finger flapping or twisting, or complex whole-body movements)
   (d) persistent preoccupation with parts of objects

B. Delays or abnormal functioning in at least one of the following areas, with onset before 3 years old: (1) social interaction, (2) language as used in social communication, or (3) symbolic or imaginative play.

C. The disturbance is not better accounted for by Rett’s Disorder or childhood disintegrative disorder.


TABLE 2  Diagnostic Criteria for 299.80: Asperger’s Disorder (Referred to as AS in This Report)

A. Qualitative impairment in social interaction, as manifested by at least two of the following:

(1) marked impairment in the use of multiple nonverbal behaviors such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction

(2) failure to develop peer relationships appropriate to developmental level

(3) a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (eg, by a lack of showing, bringing, or pointing out objects of interest to other people)

(4) lack of social or emotional reciprocity

B. Restricted repetitive and stereotyped patterns of behavior, interests, and activities, as manifested by at least one of the following:

(1) encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus

(2) apparently inflexible adherence to specific, nonfunctional routines or rituals

(3) stereotyped and repetitive motor mannerisms (eg, hand or finger flapping or twisting, or complex whole-body movements)

(4) persistent preoccupation with parts of objects

C. The disturbance causes clinically significant impairment in social, occupational, or other important areas of functioning.

D. There is no clinically significant general delay in language (eg, single words used by 2 years old, communicative phrases used by 3 years old).

E. There is no clinically significant delay in cognitive development or in the development of age-appropriate self-help skills, adaptive behavior (other than in social interaction), and curiosity about the environment in childhood.

F. Criteria are not met for another specific Pervasive Developmental Disorder or Schizophrenia.


With recent heightened public awareness, parents are more likely to raise a concern specifically about autism.35–37 In addition, as screening tools and more reliable evaluation instruments have been developed, professionals have become increasingly proficient in recognizing and diagnosing ASD. Apart from greater awareness and better ascertainment, additional reasons for the apparent increase have been debated hotly in the lay media; in fact, the publicized “autism epidemic” may be one of the most challenging public health issues today.

The prevalence of autism and, more recently, ASDs is closely linked to a history of changing criteria and diagnostic categories. Autism first appeared as a separate entity with specific criteria in the DSM-III in 1980.17 In 1987, the Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised (DSM-III-R)38 listed broadened AD criteria and the new subthreshold category of PDD-NOS, both of which promoted inclusion of milder cases. Later, these changes received criticism for being too inclusive and for promoting overdiagnosis.39 The DSM-IV criteria published in 1994 reflected the result of years of analyses to reduce the overinclusiveness of the DSM-III-R criteria; however, it included AS for the first time, which, in effect, broadened the range of disorders. Studies have revealed that the DSM-IV criteria have better specificity (0.87) than DSM-III-R criteria.40 The DSM-IV-
TR^4 criteria for AD and AS are unchanged; however, the text description of PDD-NOS was edited slightly to increase specificity. Collaboration with European groups that worked on the revised *International Statistical Classification of Diseases and Related Health Problems* (10th edition)\(^{41}\) promoted better conformity between the 2 classification systems.

AD did not become a diagnosis for which children became eligible to receive special education services until passage of the Individuals With Disabilities Education Act (IDEA) in 1990.\(^{52}\) Before the IDEA was enacted, children were labeled as having conditions such as mental retardation (MR), learning disability, speech impairment, or emotional disturbance to obtain eligibility for services.\(^{43}\) Hence, after passage of the IDEA, the resulting increase in the number of children served under the AD category reflected both newly diagnosed young children entering the school system and older children who were previously eligible for special services under a different educational label. This reflects the phenomenon of “diagnostic substitution,” whereby the number of children receiving special education under other categories (primarily MR, speech impairment, and learning disabilities) has decreased over the same time period. In addition, some increase in prevalence may be attributable to inaccuracies in diagnosis for a number of reasons, including labeling biases when schools used less rigorous criteria than those needed for a DSM diagnosis,\(^{44–48}\) when educational funding trends influenced diagnosis,\(^{49}\) and/or when parents of children with marginal criteria advocated for the AD label to qualify for supplementary services (eg, year-round schooling) described in the IDEA amendments.\(^{50,51}\) The impact of these factors on current prevalence estimates has been controversial and illustrates the reason why educational administrative data reported in some studies that receive media attention should not be considered for epidemiologic studies.\(^{47,48,52–56}\)

Just at the time when school eligibility laws were changing, the Americans With Disabilities Act of 1990\(^{57}\) was passed, obliging states to administer their programs in the most integrated settings appropriate to the needs of the person with disabilities. This was the culmination of a long series of state and federal legislation that promoted closure of institutions and encouraged governments to support families in their efforts to raise their children with disabilities at home. Thus, children with autism, especially those with comorbid MR and behavior problems who might have been institutionalized in the past, began to attend community schools and to be “counted” in educational prevalence data.

Other factors that may also be contributing to the perceived increase in prevalence include the recent identification of children with genetic disorders unrelated to ASDs who also sometimes can meet criteria for an ASD, such as Down syndrome\(^{58,59}\) and CHARGE (coloboma, heart disease, choanal atresia, retarded growth and development and/or central nervous system anomalies, genital anomalies and/or hypogonadism, and ear anomalies and/or deafness) syndrome.\(^{60}\) Finally, diagnosis of an ASD may be made in an older family member with milder symptoms that were previously unrecognized until after the diagnosis of a younger child.\(^{61}\)

Regardless of the study, the year conducted, or the reported rate of prevalence, more boys than girls are consistently found to be affected with ASDs, with male-to-female ratios ranging from 2:1 to 6.5:1.\(^{24,28,29,34,62}\) The male-to-female ratio is even higher for high-functioning autism and AS, ranging from 6:1 to as high as 15:1.\(^{63}\) (In recognition of these statistics and for the sake of brevity, this report uses masculine pronouns.)

### Etiology

ASDs are biologically based neurodevelopmental disorders that are highly heritable.\(^{64}\) Despite this fact, the exact cause still is unknown. Finding the cause has been daunting because of genetic complexity and phenotypic variation. ASDs are complex heritable disorders that involve multiple genes and demonstrate great phenotypic variation. Estimates of recurrence risks, based on family studies of idiopathic ASDs, are approximately 5% to 6% (range: 2%–8%) when there is an older sibling with an ASD and even higher when there are already 2 children with ASDs in the family.\(^{65–68}\)

In a minority of cases (<10%), ASDs may be associated with a medical condition or a known syndrome.\(^{20,21}\) Although ASDs are believed to be mainly genetic in origin, environmental factors may modulate phenotypic expression.\(^{64,69}\) Advanced paternal age\(^{70,71}\) and maternal age\(^{71,72}\) have been shown to be associated with an increased risk of having offspring with ASDs, possibly because of de novo spontaneous mutations and/or alterations in genetic imprinting. Environmental exposures may act as central nervous system teratogens in early gestational life.\(^{73}\) Some researchers have suggested that an epigenetic mechanism (heritable changes in gene expression that occur without changes in DNA sequence) may be responsible.\(^{74}\) Thus, it has become more and more apparent that the etiology is multifactorial with a variety of genetic and, to a lesser extent, environmental factors playing a role.\(^{75}\)

Two major strategies have been used in the search for the ASD genes: targeted cytogenetic/molecular studies and whole-genome screens of families of children with ASD.\(^{76–79}\) The first strategy depends on developing a hypothesis regarding the pathogenesis of ASDs, focusing on a potential candidate gene and testing it genetically for an association with ASDs. Candidate genes in ASDs include, among others, those that seem to play a role in brain development (eg, cerebellar Purkinje cell proliferation) or neurotransmitter function (eg, serotonin).\(^{80}\) The second strategy uses an indirect method and does
not require investigators to make assumptions regarding the mechanism of inheritance. Instead, families with multiple members who demonstrate an ASD (multiplex families) are studied to identify recurring DNA markers (break points, translocations, duplications, and deletions) present in affected members but not in unaffected members. Unfortunately, progress in determining a genetic etiology using this method has been impaired, because the phenotypic end points of ASDs are not well defined. Changing DSM criteria and inconsistent ascertainment strategies, which results in a hazy delineation between affected versus unaffected family members, obscure outcomes and challenge interpretation of results.87 This phenotypic heterogeneity has challenged molecular searches for the ASD gene(s) despite several genomewide screens of the International Molecular Genetic Study of Autism Consortium and multicenter collaborative efforts over the past couple of decades.78,81–84 Although at least 1 autism-linked abnormality has been found on almost every chromosome, sites on a few chromosomes (X, 2, 3, 7, 15, 17, and 22) seem to be more promising than others.57,68,75,79,85–90 Maternally derived 15q duplications are common; depending on the investigator, yields vary from 1% to 10%,91 with most in the range of 1% to 3%.92,93 Patients with these duplications may not display dysmorphic features, but they often have hypotonia and/or global developmental delay (GDD) and may develop seizures later. The abnormality can often be identified on high-resolution karyotype analysis. Other less common abnormalities have also been reported.94

Finally, the male predominance noted above also suggests a genetic role in the inheritance of autism. Several genetic processes can lead to male predominance, including causative genes located on the X chromosome (X-linked disorders) and imprinted genes, but the reason for male predominance in autism is not completely understood.95

In a discussion of etiology, subtyping ASDs as either idiopathic or secondary is helpful.57,79,95 For the purposes of this discussion, the term “idiopathic” ASDs refers to cases in which children meet criteria for ASDs but do not have a comorbid associated medical condition known to cause ASDs. Most individuals with an ASD have the idiopathic type. Children with idiopathic ASDs demonstrate variable behavioral phenotypes, are somewhat less likely to have comorbid GDD/MR, and generally do not have dysmorphic features that herald a recognizable syndrome. Nevertheless, twin and family studies have revealed that idiopathic ASDs are heritable and have a recurrence rate of 5% to 6%.97,94,95 The term “secondary” ASDs refers to cases with an identifiable syndrome or medical disorder known to be associated with ASDs. Whereas earlier reviews reported that the proportion of individuals with ASDs who have a comorbid syndrome or medical condition was 10% to 20%,5,96–98 the proportion has decreased to less than 10% when using more recent data sets.79,89,95–99 In a meta-analysis of 23 epidemiologic studies, Chakrabarti and Fombonne80,21 revealed that a recognizable condition was identified in only 6% of those with a confirmed ASD. The rate of coexisting MR (cognitive impairment associated with an IQ of <70) in children with ASDs seemed to decrease from 90% before the 1990s to less than 50% after 2000,28,29,34,35,102,103 possibly because of improved methods in testing intelligence in this population and to the increased awareness of children with ASD with milder features and higher functioning. This trend is important, because coexisting severe MR, especially in the presence of dysmorphic features, increases the likelihood of identifying a known disorder.89,104–108 Neurogenetic syndromes that seem to play a causative role or otherwise are associated with ASDs include, but are not limited to:

- Fragile X syndrome109,110: Fragile X syndrome is the most common known genetic cause of AD and of MR in males. The phenotype includes MR, macrocephaly, large pinnae, large testicles (particularly after puberty), hypotonia, and joint hyperextensibility. Identifying a patient with fragile X syndrome is important for genetic counseling purposes, because the diagnosis has implications for other family members. Depending on the prevalence of comorbid MR in study subjects with ASD, the etiologic yield of fragile X syndrome–DNA testing has ranged from 0% to 8%, with a median of approximately 3% to 4%.99,109,111 On the other hand, as many as 30% to 50% of individuals with genetically confirmed fragile X syndrome will demonstrate some characteristics of ASDs.102,110

- Neurocutaneous disorders: Tuberous sclerosis112–116 is characterized by hypopigmented macules (sometimes requiring a Wood’s lamp examination for visualization in young children), fibroangioma, kidney lesions, central nervous system hamartomas, seizures, MR, and autistic and/or attention-deficit/hyperactivity disorder (ADHD)–like behaviors. Although tuberous sclerosis is a dominant disorder (with genes located at 9q and 16p), most cases represent new mutations. Although it is the most common neurocutaneous disorder, neurofibromatosis is less likely to be associated with ASDs. It also is autosomal-dominant, with half of cases representing new mutations of the neurofibromatosis 1 gene on 17q.117 It is characterized by café au lait macules and freckling in the axillary and inguinal regions, neurofibromas, and ocular Lisch nodules. Although most patients have a benign course and normal intelligence, a small subset of individuals have MR and behavioral features that are consistent with ASDs.

- Phenylketonuria118: phenylketonuria now is a rare cause of ASDs and MR in the United States, because it
is preventable as a result of newborn screening and dietary intervention.

- **Fetal alcohol syndrome**: Children who are exposed to alcohol during gestation have an increased risk of ASDs in addition to other neurodevelopmental disorders.

- **Angelman syndrome**: Angelman syndrome is associated with loss of the maternally expressed ubiquitin-protein ligase gene (UBE3A) on 15q through deletion, paternal uniparental disomy, or imprinting errors. Children with Angelman syndrome present with GDD (and often are nonverbal), hypotonia in early childhood, wide-based ataxic gait, seizures, and progressive spasticity. Angelman syndrome associated with a deletion of 15q can be detected with fluorescence in situ hybridization (FISH) testing; however, when it results from uniparental disomy, methylation studies are necessary.

- **Rett syndrome**: Rett syndrome usually presents with a classic phenotype and should be considered in all females who demonstrate autistic-like regression, especially if they have microcephaly, seizures, and hand-wringing stereotypies. Retrospective videos have revealed early subtle motor symptoms during the first year of life. Now that it is possible to confirm this diagnosis with DNA testing (methyl CpG-binding protein 2 [MECP2]) in approximately 80% of cases, it has become apparent that there is a spectrum of severity, and some patients may present with atypical features including those consistent with ASDs. Rett syndrome is much less common in males, and the presentation is more varied. Some males die in infancy as a result of neonatal encephalopathy; others with comorbid Klinefelter syndrome (as well as a few males [in isolated case reports] with a normal number of sex chromosomes) demonstrate more classic symptoms.

- **Smith-Lemli-Opitz syndrome**: Smith-Lemli-Opitz syndrome is a rare (1 in 20 000) autosomal-recessive disorder caused by a metabolic error in cholesterol biosynthesis. Although most patients present with multiple congenital anomalies, failure to thrive, and MR, some may present with subtle physical features such as webbing (syndactyly) of the second and third toes, mild hypotonia, and autistic features. Recurrence risk is 25%; thus, appropriate genetic counseling is important.

Whether the aforementioned conditions play a direct or indirect etiologic role or simply are associated with ASDs, they still represent a small minority of patients with ASDs. Conversely, a few children with genetic syndromes that are characterized by features quite different from ASDs also may meet DSM-IV-TR criteria. For example, recent studies have reported that 6% to 7% of children with Down syndrome (typically characterized by relatively good social skills compared with those in other domains) and almost 50% of children with CHARGE syndrome (associated with mutations of the CHD7 gene) meet criteria for one of the ASDs. There have also been a few isolated reports of a mitochondrial and/or metabolic abnormality (eg, carnitine deficiency) being associated with an ASD, but the significance of these reports is not clear.

Increased and decreased levels of T lymphocytes, immunoglobulins, and antibrain autoantibodies in the systemic circulation have been reported. These have been observed chiefly in retrospective case studies of patients with idiopathic ASDs, but systematic prospective studies have confirmed neither their existence nor their relevance. Prospective studies have revealed that, except for a few individuals with recurrent infections, healthy children with ASDs generally have normal immune function. Some studies have reported increased rates of autoimmune disorders in families of children with ASDs, particularly in the mothers (eg, thyroid disorders and psoriasis); however, the relevance of these common disorders to ASDs in children is unknown. Furthermore, studies have shown no increase in autoimmune disorders of the central nervous system, and patients with ASDs did not themselves exhibit autoimmune disorders. The contribution of possible immunologic dysfunction remains to be further defined.

**Environmental Issues**

Regardless of the mechanism, a review of studies published in the past 50 years revealed convincing evidence that most cases of ASDs result from interacting genetic factors. However, the expression of the autism gene(s) may be influenced by environmental factors. Although currently under investigation, these factors may represent a “second-hit” phenomenon that primarily occurs during fetal brain development. That is, environmental factors may modulate already existing genetic factors responsible for the manifestation of ASDs in individual children.

**Prenatal Period**

Because many of the developmental brain abnormalities known to be associated with ASDs occur during the first and second trimesters of pregnancy, environmental factors (eg, teratogens, such as thalidomide and valproic acid) are more likely to play a role in the fetus via maternal factors. It is possible that maternal illness (eg, rubella) during pregnancy plays a role. Recently, the possible association between fetal testosterone concentration and certain autistic behaviors such as abnormal social relationships and restricted interests at 4 years of age was investigated.
Perinatal Period

The effects of birth weight, duration of gestation, and events around the time of birth have been investigated also, but findings have not been consistent. A significant association between term newborn encephalopathy and children later diagnosed with ASD was reported recently. Badawi et al reported that 5% of survivors of newborn encephalopathy were diagnosed with an ASD, which represented an almost sixfold increase compared with matched controls. This increase may represent a genetically derived predisposition (which makes the infants vulnerable to both encephalopathy and ASD) or an independent mechanism.

Postnatal Period

Etiologic possibilities occurring after birth have been proposed—in particular, measles-mumps-rubella (MMR) vaccine and mercury-containing vaccines. In 2001, the Institute of Medicine reviewed epidemiologic population-based studies and concluded that there was no evidence of a causal association between the MMR vaccine and autism. Studies that examined the association between MMR vaccine and autism since the publication of that review have supported this conclusion. Questions also have been raised about the effects of environmental mercury exposure (including mercury-containing vaccines) on brain development in ASDs and other developmental disabilities. Mercury, in its organic form, is a known neurotoxin with neurologic sequelae, including motor impairment and visual and intellectual deficits, depending on the age at exposure and the type of mercury. There is no evidence to date that children with neurodevelopmental disabilities, including autism, in the United States have increased mercury concentrations or environmental exposures. Using large data sets from the United States, Sweden, and Denmark, to date, no consistent association has been found between thimerosal-containing vaccines and neurodevelopmental outcomes or prevalence of ASDs. Despite evidence to the contrary, a recent survey of parents of children with ASDs revealed that 54% believed that their child’s ASD was caused by immunizations; 53% thought it was caused by genetics.

Although the previous discussion reveals the wide variety of conditions known to be associated with ASDs, currently, an etiologic investigation of the individual child with an ASD infrequently identifies a known cause in the absence of GDD/MR, dysmorphic features, a positive family history, and/or a local neurologic examination.

NEUROPATHOLOGY AND NEUROIMAGING

In recent years, intense research efforts have focused on elucidating the neurobiological basis of ASDs. A growing body of evidence from neuropathology and neuroimaging studies indicates that there are fundamental differences in brain growth and organization in people with ASDs that have their origin in the prenatal period but extend through early childhood and into adulthood.

Neuropathologic studies of brain tissue from people with autism have revealed several abnormalities including:

- reduced numbers of Purkinje cells in the cerebellum;
- abnormal maturation of the forebrain limbic system, including reduced neuronal size, increased cell-packing density, and decreased complexity of the neuropil (ie, the complex net of axonal, dendritic, and glial branching in which the nerve cell is embedded);
- abnormalities in frontal and temporal lobe cortical minicolumns, which are more numerous, smaller, and less compact in their cellular configuration and demonstrate reduced neuropil space in the periphery;
- developmental changes in cell size and number in the nucleus of the diagonal band of Broca, deep cerebellar nuclei, and inferior olive; and
- brainstem abnormalities and neocortical malformations (eg, heterotopias).

The most consistent neuropathologic findings suggest pathology that arises in utero. The association of increased risk of ASDs associated with prenatal exposure to teratogens, such as thalidomide and valproic acid, suggests that early insults during critical periods of brain development (as early as 20–24 days after conception in the case of thalidomide) may be sufficient to cause ASDs. However, all of these neuropathologic findings are based on detailed study of a relatively small number of brains, and further investigation is required. Limited availability of brain tissue from people with well-characterized ASDs and age-matched controls has impeded neuropathologic investigations. Efforts to remedy this are underway with the establishment of the Autism Tissue Project (1-800-272-4622 [for physicians] or 1-877-333-0999 [for families]; www.memoriesofhope.org).

Kanner, in his initial clinical description of autism, noted large head size in several of his patients. Increased head circumference has since been shown to be a common physical finding in children with ASDs, and 20% to 30% have macrocephaly, defined as a head circumference that measures more than 2 SDs above the mean. MRI studies have supported the finding of increased brain volume in children with ASDs, with 90% of toddlers with ASDs having larger-than-normal brain volumes in 1 study. Postmortem brain weights also are increased. Children later diagnosed with an ASD have been shown, as a group, to have average or below-average head circumference at birth, with acceleration in brain growth during the first year of life.
life, leading to above-average head circumference or overt macrocephaly. Fewer adults with ASDs have been found to exhibit increased brain size compared with controls, indicating that there may be deceleration of brain growth at some point beyond early childhood. It is interesting to note that increased blood concentrations of brain-derived neurotrophic factor and several other neurotrophins have been detected in newborn infants who are later diagnosed with ASDs. This finding, if replicated, may have implications regarding the mechanism of early brain overgrowth. Age-related differences in serotonin synthesis capacity also have been demonstrated between children with ASDs and children in control groups, which leads to speculation regarding the neurotrophic role of serotonin in abnormal brain growth and organization in children with ASDs.

In addition to whole-brain volume differences, specific regional gray- and white-matter volumetric differences have been described. The frontal, limbic, basal ganglia, and cerebellar regions have been implicated most consistently. Abnormalities in sulcal and gyral anatomy have been found by using surface-mapping techniques. The regional gray- and white-matter volume differences also seem to be age related, although larger cross-sectional studies and longitudinal studies are needed to clarify the meaning of these findings.

A variety of functional MRI studies during cognitive tasks or in response to visual or auditory stimuli suggest that individuals with ASDs use different cognitive strategies and, in some cases, different brain areas to process certain types of information. For example, functional neuroimaging techniques have indicated the presence of abnormalities in face recognition and executive functioning in adults with high-functioning ASDs. Hypoactivation of the fusiform gyrus in face-recognition tasks has been one of the most consistent findings and, in concert with abnormalities in amygdala activation, may relate to the abnormalities in gaze fixation that are seen in people with ASDs. Functional MRI evidence has also been used to postulate impaired “connectivity” between various cortical regions in the brains of people with ASDs. Most recently, some investigators have attempted to explain deficits in empathy, imitation, and language as abnormalities in the functioning of mirror neuron systems. These systems are a newly discovered subset of cells found in several areas of the brain that seem to fire when an individual simply observes another’s actions—that is, it seems they directly reflect actions performed by another in the observer’s brain. They also may play a role in the ability to recognize and empathize with or “mirror” the feelings of others. These functional brain differences provide intriguing links between the neuroanatomical substrate and the characteristic clinical features of people with ASDs.

Although neuroimaging research has identified volumetric and other abnormalities in groups of patients with ASDs compared with controls, a reliable marker has not been identified, and routine clinical neuroimaging for individuals with ASDs is not recommended.

**CLINICAL SIGNS**

Whereas severe social skills deficits and restricted, repetitive, and stereotyped patterns of behavior, interests, and activities are core features of all ASDs, significant language delays are characteristic of only AD and PDD-NOS. One of the most challenging aspects in recognizing ASDs is the wide heterogeneity of features in individual children. There is no pathognomonic feature; however, a few of the early social deficits (eg, delayed or absent joint attention) seem to be fairly reliable red flags for ASDs. The autism spectrum encompasses an extremely heterogeneous phenotype with indistinct end points, especially at the mild end of the spectrum. The severity of each of the core deficits varies significantly among children with ASDs.

Although the social deficits occur earlier and may be more specific, they can be subtle and less often recognized or articulated by parents. Speech delays usually prompt parents to raise concerns to their child’s PCP. Most parents become concerned between 15 and 18 months of age but may delay discussing their concerns with their child’s physician for several months. Recently, the media and public agencies have raised public awareness about the importance of recognizing the early signs, including those present during the first years of life. This being the case, it is anticipated that parents may begin to voice concerns to their infant’s pediatrician earlier and that these concerns may now target the often earlier-appearing social deficits. Presentations can differ widely from one child to the next; some are perceived by parents as “different” during the first few months of life, others present with delayed speech development during the second year of life, and still others may appear to be normal only to regress and lose skills after the first year of life. AS in children may go unnoticed until they are of school age, when teachers notice difficulties with peer interactions. Expanded reviews regarding early signs are available.

**Social Skills Deficits**

Although more specific than language deficits, social deficits appearing in the first 2 years of life often have escaped parent recognition. Children with ASDs universally demonstrate deficits in social relatedness defined as the inherent drive to connect with others and share complementary feeling states. Children with ASDs often do not appear to seek connectedness; they are content being alone, ignore their parents’ bids for attention, and seldom make eye contact or bid for others’ attention with gestures or vocalizations. In later
years, they have difficulty sharing the emotional state of others in cooperative games and group settings and may have few, if any, friends.

Deficits in JA seem to be one of the most distinguishing characteristics of very young children with ASDs. JA is a normal, spontaneously occurring behavior whereby the infant shows enjoyment in sharing an object (or event) with another person by looking back and forth between the two. Later, gestures and/or speech also can be used to engage another's attention with regard to the objects and events simply for the enjoyment of sharing the experiences. Just like other developmental skills, development of JA skills is stepwise; it occurs in stages beginning in the first few months of life. Similar to language skills, receptive JA skills usually are mastered before expressive ones. JA begins with joyous smiling in recognition of and response to a parent or familiar caregiver's smiles and vocalizations. At approximately 8 months of age, an infant will follow the parent's gaze and look in the same direction when a parent looks away (i.e., to check the time). Children begin to “follow a point” at approximately 10 to 12 months of age. If a parent points in the direction of an interesting object or event and says, “Look!” the typically developing child will look in the intended direction and then, after seeing the object/event, look back at the parent in acknowledgment and shared expression. Infants with ASD may not follow a point, even when one tries repeatedly in a loud voice calling their name or uses physical prompts, such as touching the child's shoulder before pointing. They may look in the indicated direction eventually, but this is not followed by shared looking and expression.

At approximately 12 to 14 months of age, the typically developing child will begin himself to initiate a point, at first to request a desired object that is out of reach and, a couple of months later, to draw the parent's hand to lead him or her to the object. At 14 to 16 months of age, the typically developing child will begin to point simply to “comment” about or “share” an interesting object/event (which is called “protodeclarative pointing”). As he points, he will look alternatively to the object/event of interest and the parent. It is the shared social experience, not the tangible object/event, that the child seeks. Children with ASDs consistently fail to point to “comment” at age-appropriate times, and when they do, they are less likely to show positive affect and connectedness during the act. Some high-functioning children with ASDs may point to label objects, shapes, and colors that they have learned in a rote fashion, but this often is done without any intent of communicating in a social context and is not considered JA. Mastery of JA seems to be necessary for functional language development; in fact, mastery of protodeclarative pointing seems to be a reliable predictor of functional language development within 1 year. JA skills progress to involve ongoing back-and-forth bids for attention and social interactions with multiple emotional expressions, sounds, words, and other gestures.

Orienting to social stimuli—in particular, turning consistently to respond to one’s own name—is an early skill (8–10 months of age) that often is deficient in children with ASDs. However, it is not specific to children with ASDs, because children with hearing impairments also may fail to orient to their name. In fact, parents of children later diagnosed with ASDs often raise a concern about hearing. Hearing seems “selective” in that children with ASDs may hear and attend well to environmental sounds but not to human voices. Social referencing is the ability to recognize the emotional states of others as they respond to various stimuli. When faced with a novel situation, a typically developing infant might look to his mother for an indication of delight, anger, or fear in her facial expression. His facial expression then usually will mimic hers, although he may not fully understand the situation. A child with an ASD engages in less imitation.

Because children with ASDs lack fundamental social skill building blocks, they may be less likely to develop appropriate peer relationships according to age and language ability. They may have few or no friends, and when they do, the relationships may evolve around the child’s own special interests. Another factor that impedes lasting friendships is impaired central coherence or the inability to interpret stimuli in a global way. Instead, they focus on the parts, make less use of context, and miss the “big picture,” which makes social interactions challenging. They also have difficulties understanding the perspective of others or lack “theory-of-mind” (ToM) skills. ToM is the awareness that others have thoughts and emotions that are independent from one’s own; it is the ability that allows one to infer states of mind on the basis of external behavior. Typically developing children begin to have some sense of mental states of others by 4 years of age. Because of ToM impairments, children with ASDs have difficulties with empathy, sharing, and comforting. Baron-Cohen coined the term “mindblindness” when referring to persons with ASDs who demonstrate severe ToM deficits.

Communication Deficits

Most children who are later diagnosed with AD and PPD-NOS present to their PCP with “speech delay,” al-
Echolalia, sometimes called “parroting,” is the repetition of another person’s speech. Echolalia is classified as “immediate” when the child repeats vocalizations promptly after hearing them or “delayed” when there is a time lapse (hours, days, weeks). Typically developing children pass through a “vocabulary-burst stage,” when brief periods of immediate echolalia are not unusual.197 On the other hand, echolalia in children with ASDs may persist throughout the life span and consist of a mixture of immediate and delayed varieties. Utterances of children with ASDs may be more clearly articulated, have a more monotone quality, and/or consist of larger verbal “chunks” (ie, entire television advertisement jingles, video reenactments, or recitations of nursery rhymes) than those of typically developing children. Sometimes, echolalia may even give the impression of “advanced” speech because of sophisticated vocabulary, grammar, and syntax. The clinician should be careful to differentiate between typical and autistic echolalia; usually, a formal evaluation by a speech-language pathologist (SLP) is needed. Such an assessment also may reveal a dissociation between these “advanced” expressive skills and delayed receptive ones in that the child may be unable to follow simple 1-step commands, which is a 12- to 14-month-old skill. Some parents will note that their child seems overly “independent” because, rather than ask for desired objects, he uses advanced motor skills to obtain them himself (ie, moving a stool to a counter top to obtain an object at an age younger than typically expected). Some children with ASDs become quite skilled at rote labeling colors, shapes, numbers, and letters of the alphabet, yet they are unable to point to them when asked to do so by another or incorporate the labels into functional language. A few may later develop hyperlexia or advanced verbal reading without corresponding comprehension skills.

Some children with ASDs say “pop-up words” without any apparent stimulus or communicative intent. They are spontaneous and inconsistent, although sometimes they may occur during acutely stressful situations. These words are said out of context for a short period of time (days or weeks) and then, as suddenly as they might pop up for no apparent reason, they disappear.197,119 Children with ASDs also may develop “language” in overlearned or gestalt phrases that are acquired and spoken almost as a single “giant-word” (ie, What is it? I don’t know). At the same time, they are unable to combine words in novel or original phrases or sentences that convey true meaning.

Although lack of speech, scripted speech, parroting without communicative intent, and pop-up and giant words are common classic presentations, earlier prespeech deficits often exist that, if detected, could facilitate earlier diagnosis.* These deficits include:

- lack of appropriate gaze;
- lack of warm, joyful expressions with gaze;
- lack of the alternating to-and-fro pattern of vocalizations between infant and parent that usually occurs at approximately 6 months of age (ie, infants with ASDs usually continue vocalizing without regard for the parent’s speech);
- lack of recognition of mother’s (or father’s or consistent caregiver’s) voice;
- disregard for vocalizations (ie, lack of response to name), yet keen awareness for environmental sounds;
- delayed onset of babbling past 9 months of age;
- decreased or absent use of prespeech gestures (waving, pointing, showing);
- lack of expressions such as “oh oh” or “huh”;
- lack of interest or response of any kind to neutral statements (eg, “Oh no, it’s raining again!”)

The AAP brochure “Is Your One-Year-Old Communicating With You?”13 was developed to help raise parent and physician awareness of these earlier social communication milestones and to promote recognition of symptoms of ASDs before 18 months of age.

Regression

Approximately 25% to 30% of children with ASDs begin to say words but then stop speaking, often between the ages of 15 and 24 months.199,200,208 Regression of skills in children with ASDs may also include loss of gestural communication (wave, point, etc) and social skills (eg, eye contact and response to praise) or a combination of both. Regression can be gradual or sudden, and it may be superimposed on subtle preexisting developmental delays or atypical development, such as an unusually intense interest in objects or other nonsocial stimuli during the first year of life.205 Although it may be tempting to attribute regression to environmental stressors (eg, birth of a new sibling or a move to a new house), this results in a delay in diagnosis. Regression is a well-documented hallmark of ASDs and should always alert the PCP to consider ASDs.

*Refs 197, 204, 213, 214, 219, and 222.
Asperger Syndrome

Children with AS may have mild or limited speech delays (see the DSM-IV-TR criteria in Table 2) and escape recognition until preschool or early school age, when their inability to make friends becomes a concern. Although often unnoticed, language development usually is atypical. Children with AS often are quite verbal about a certain topic of interest, but they are unable to express simple feelings or recognize the feelings and viewpoints of others. Speech may be fluent but limited to only a few topics, typically those that hold a strong, all-consuming interest for the child. Speech also can be overly formal (pedantic), which is a reason why children with AS sometimes are described as “little professors.” Children with AS also have deficits in the social use of language (pragmatics): how to choose a topic of conversation; understanding and producing appropriate tempo, facial expression, and body language during conversation; turn taking; recognizing when the partner has lost interest in a topic; knowing when to start, sustain, and end a conversation on the basis of listener cues; knowing when and how to repair a communication breakdown; and using the appropriate degree of formality and politeness. Children with AS especially have difficulty sustaining a conversation on a topic that is initiated by another. Language may seem odd, self-centered, and not listener responsive and results in a monotone monologue. They may demonstrate unique delivery of speech (prosody) in regard to intonation, volume, rhythm, pitch, and personal space that also tends to disregard listener needs. Children with AS may have difficulty with abstract reasoning and discussion of thoughts and opinions of others. Inability to discern and judge the conversational intents of others, especially when their conversation includes words or phrases with ambiguous meanings, impairs their ability to understand metaphors, humor, teasing idioms, irony, lies, jokes, and faux pas.

Older children with high-functioning AD or PDD-NOS and fluent speech also may demonstrate some of the above-mentioned language characteristics.

Play Skills

Lack of, or significantly delayed, pretend play skills coupled with persistent sensory-motor and/or ritualistic play are characteristic of ASDs. Some children with severe ASDs may never progress past the sensory-motor play stage. They mouth, twirl, bang, and manipulate objects in a stereotypic or ritualistic manner. The play of children with ASDs often is repetitive and lacks creativity and imitation. Typical examples include spinning the wheels or lining up cars instead of “driving” them, arranging crayons instead of coloring with them, or stacking blocks in the same sequence time after time. Often they prefer to play with common objects (string, sticks, rocks, or ballpoint pens) rather than store-bought toys with the exception of trains or characters from favorite videos and television shows. Puzzles, especially shape-matching ones and computerized “puzzle games,” also are quite popular. Children with ASDs often are content to play alone for hours, requiring little attention or supervision. Often this “play” is either constructive (puzzles, computer games, and blocks), ritualistic (lining objects up or sorting/matching shapes or colors) or sensory-motor (mouthing, banging, twirling) in nature. Children with ASDs may seem to enjoy chase games and roughhousing, but it is often the sensory-motor aspects of these activities, rather than their social aspects, that are enjoyable. They have trouble interacting in groups and cooperating in the social rules of more sophisticated games. Often they are left out, ignored, and at high risk of being victimized and bullied by peers.

Restricted, Repetitive, and Stereotyped Patterns of Behavior, Interests, and Activities

Children with ASDs can demonstrate atypical behaviors in a variety of areas including peculiar mannerisms, unusual attachments to objects, obsessions, compulsions, self-injurious behaviors, and stereotypies. Stereotypies are repetitive, nonfunctional, atypical behaviors such as hand flapping, finger movements, rocking, or twirling. Although most stereotypies are harmless, they are problematic in that they may prevent the child from accomplishing a task or learning new skills. Although stereotypies are distinctive and obvious, they are not specific to children with ASDs, because many children with profound MR and/or severe sensory deficits also demonstrate stereotypies. Even typically developing toddlers, especially before the onset of fluent language, may flap their arms briefly when they are excited or frustrated. Stereotypies associated with ASDs often do not appear until after 3 years of age and commonly manifest as finger flicking, unusual eye gazing, habitual toe walking, and/or persistent sniffing and licking of nonfood items.

Although most children, at some time during their early development, form attachments with a stuffed animal, special pillow, or blanket, children with ASDs may prefer hard items (ballpoint pens, flashlight, keys, action figures, etc). Moreover, the attachment is more persistent, in that they may insist on holding the object at all times, although these are rarely, if at all, used in real “play.” Whereas younger children with ASDs may have restricted interests in regards to objects, the restricted interests in those with AS more often relate to topics and facts. For example, rather than carrying a toy train at all times, there is an obsession with train schedules. Sometimes the item/topic of interest may be typical for any child, but it is the degree of interest that is abnormal. For example, similar to typically developing children, a child with an ASD may be fascinated with dinosaurs, but he knows far more details about them and persists in playing or discussing them to the exclusion of all else.
Perseveration, or continuation of speech or play to an exceptional degree or beyond a desired point, is common in children with ASDs. Children with ASDs may protest vigorously when forced to transition from an activity or topic of interest or when a usual routine is changed. Without warning, these protests may quickly escalate to severe and prolonged temper tantrums characterized by aggression or self-injurious behaviors.

Self-injurious behaviors (head banging, skin picking, eye poking, hand biting) are stereotypies that may cause bodily harm and are more common in children with severe GDD/MR (intellectual disabilities) or ASDs with comorbid GDD/MR. Self-injurious behaviors may be precipitated by frustration during unsuccessful communication attempts, transitions, anxiety in new environments, boredom, depression, fatigue, sleep deprivation, or pain. The presence of self-injurious behaviors, aggression, and other extreme behaviors may prevent the child from participating in integrated activities in the community with typically developing peers and cause significant family stress.

Additional Coexisting Conditions That Are Not Core Features in the DSM-IV-TR

Cognitive Abnormalities (GDD/MR or Intellectual Disability, Learning Differences, and Splinter/Savant Skills)

The prevalence of comorbid GDD/MR or intellectual disability (the appropriate term depends on age and availability of both a standardized IQ score and a formal assessment of adaptive skills) with ASDs was estimated to be approximately 90% before 1990. On the basis of the later studies published in the 1990s, consensus guidelines reported the prevalence as approximately 70% to 75%. Prevalence studies published in the new millennium have reported rates of ASDs with comorbid GDD/MR of just under 50%, whereas 2 English studies reported rates as low as 26% to 29%. Better ascertainment of children without cognitive deficits (in particular, AS, which by definition is characterized by normal intelligence), improved professional training, and more effective strategies/tools for evaluating cognitive abilities in children with ASDs all may contribute to the decreasing prevalence of comorbid GDD/MR.

One unique characteristic of ASDs is the “unevenness” of skills. Abilities may be significantly delayed in some areas of development yet “advanced” in others, often because of exceptional focusing, memory, calculation, music, or art abilities. They may be labeled as “splinter skills” when they serve no purpose in day-to-day life and do not improve functional outcomes. Rarely, highly developed talents or savant skills may promote a vocation that provides financial independence and, occasionally, national recognition.

Sensory-Motor Symptoms

Although sensory symptoms (eg, hyperacusis) are more frequent and prominent in children with ASDs, there is no evidence that sensory symptoms differentiate children with ASDs from children with other developmental disabilities. Children with ASDs may demonstrate simultaneous hyposensitivities and hypersensitivities for stimuli within the same sensory modality. For example, they may seem overly sensitive to certain environmental noises but lack response to human voice, or they may visually inspect the details of an object but not notice the comings and goings of other people in the room. Others may have oral aversions and/or total-body “tactile defensiveness” to soft touch (fabric bumps on socks and sweatshirts) or hugs yet be insensitive to pain. Sensory factors related to food, such as texture, color, and taste, may lead to highly restricted diets. More research is needed to operationalize the concept of sensory integration and possible interventions and define its role in ASDs.

In addition to unusual motor stereotypies that serve as defining characteristics of ASDs discussed previously, some children with ASDs also may demonstrate atypical motor development, poor coordination, or deficits in praxis (motor planning, execution, and sequencing). Some investigators believe that, although not a defining characteristic by DSM standards, motor clumsiness is a distinguishing characteristic of AS. Finally, some children may appear to be “hyperactive” and motor driven with an exterior focus of attention and actually meet criteria for comorbid ADHD (although current DSM-IV-TR criteria exclude making the diagnosis of ADHD in the presence of an ASD). Other children may be hypoactive and withdrawn and have an interior focus of attention.

In summary, ASDs are characterized by a broad array of clinical features; some are more specific to ASDs than others. Familiarity with the early social and preverbal communication deficits will help the PCP recognize ASDs earlier, which should, in turn, facilitate the prompt initiation of appropriate interventions.

SURVEILLANCE AND SCREENING

Because the prevalence of ASDs is approximately 6 to 7 per 1000 in the United States, PCPs are likely to provide care for children with ASDs. Early identification of ASDs is important, because it allows early intervention, etiologic investigation, and counseling regarding recurrence risk. The medical home is an important setting for surveillance and screening to detect ASDs and other developmental disorders. In the past, it was not unusual for parents’ initial concerns to be dismissed and for diagnosis and intervention to be delayed. In a recent study in metropolitan Atlanta, Georgia, the mean age of the first evaluation for 115 8-year-old chil-
Children with ASDs was 48 months, and the mean age of the first ASD diagnosis was 61 months. The goal of this clinical report is to help pediatricians identify children at an earlier age who are at risk of an ASD. An ASD-specific surveillance and screening algorithm (Fig 1) has been developed to facilitate the identification process. It builds on the developmental surveillance and screening algorithm for pediatric preventive care visits that was published in the 2006 policy statement “Identifying Infants and Young Children With Developmental Disorders in the Medical Home: An Algorithm for Developmental Surveillance and Screening.”

**General Developmental Surveillance and Screening**

According to the AAP policy statement “Identifying Infants and Young Children With Developmental Disorders in the Medical Home: An Algorithm for Developmental Surveillance and Screening,” surveillance is the ongoing process of identifying children who may be at risk of developmental delays, and “screening” is the use of standardized tools at specific intervals to support and refine the risk. As an analogy, whereas surveillance represents a “moving picture” of the child’s unfolding development, screening represents “snapshots” of the child’s development at specific times. Developmental surveillance should occur at every preventive visit throughout childhood and includes the following components: eliciting and attending to the parents’ concerns; maintaining a developmental history; making accurate and informed observations of the child; identifying the presence of risk and protective factors; and documenting the process and findings. Research has revealed that parents have valid concerns about their children’s development, although careful interpretation of the concerns is needed. However, parental concerns may not be shared if the PCP does not ask about the child’s development, and lack of parental concern about development does not imply typical development. Therefore, a systematic surveillance strategy must be used for all children. Screening with a standardized developmental tool should be performed whenever concerns are raised through the ongoing surveillance process. The AAP also recommends that all children be screened with a standardized developmental tool at specific intervals (ie, at the 9-, 18-, and 24- or 30-month visits) regardless of whether a concern has been raised or a risk has been identified during the surveillance process (see the AAP developmental screening and surveillance algorithm).

**Surveillance for ASD**

Surveillance at the first preventive care visit (Fig 1, Steps 1a and 2) should begin with a family history to determine if there are any family members, especially a sibling, who have been diagnosed with ASDs. Because the risk of having symptoms of ASDs in younger siblings of children with ASDs is approximately 10 times higher, the pediatrician needs to be extra vigilant in monitoring for early abnormal signs. Studies of infant siblings with ASDs have revealed that very subtle early signs do exist and can be perceived during the first year of life. Until recently, most knowledge regarding very early signs was obtained from retrospective systematic reviews of home videos, particularly first birthday party videos. Studies of home videos at earlier ages have provided additional retrospective information that reveals subtle abnormalities in infants who were thought to be typically developing and later diagnosed as having regressive autism. Several groups of investigators are following younger siblings of children diagnosed with ASDs and providing prospective information as symptoms emerge in these infants at high risk. Preliminary results support the feasibility of recognizing subtle signs of ASDs in infants at high risk. Some of the very early signs reported by several investigators include extremes of temperament and behavior (ranging from marked irritability to alarming passivity); poor eye contact; poor response to other’s voices, especially to one’s name being called; poor attempts at interactive play; more interest in looking at objects than people; delayed pointing to request or share; decreased to-and-fro babbling and jargoning; and lack of warm, joyful, reciprocating expressions.

Surveillance should include asking parents open-ended questions about their concerns regarding the child’s development and behavior (Step 2). Parental concerns about inconsistent hearing or unusual responsiveness also are important; for example, parents may notice that the child responds consistently to a quiet sound, such as the crinkle of a plastic snack bag, but not to a human voice calling his name. In addition, parent concerns may be stimulated by comments made by other care providers such as child care staff or preschool teachers. Recently, however, the public media have significantly increased awareness of ASDs and sometimes has stimulated unnecessary concerns. The AAP patient education brochure “Is Your One-Year-Old Communicating With You?” can be distributed to all parents at their child’s 9- or 12-month preventive visit to educate them about early social communication milestones to help them identify valid areas of concern.

Surveillance also includes asking age-specific questions about whether certain developmental milestones have been attained. When this approach is used, it is important to include social and emotional milestones in addition to the traditional motor, language, and problem-solving milestones (see www.firstsigns.org). To recognize ASDs as early as possible, it is important to ask about the development of verbal and nonverbal communication, reciprocal social interaction (including eye contact, JA and social referencing, and sharing of interests or achievements), and representational or pretend play skills. The American Academy of Neurology and...
Surveillance and Screening Algorithm: Autism Spectrum Disorders (ASDs)

1a: Pediatric Patient at Preventive Care Visit

1b: Extra Visit for Autism-Related Concern, ASD Risk Factor, or Other Developmental/Behavioral Concern

2: Perform Surveillance
   Score 1 for Each Risk Factor:
   - Sibling with ASD
   - Parental Concern
   - Other Caregiver Concern
   - Pediatrician Concern

3: What is the Score?
   Score = 2+
   Score = 1
   Score = 0

3a: Is the Patient at Least 18-Months Old?
   Yes
   No

5a: Evaluate Social-Communication Skills

5b: Administer ASD-Specific Screening Tool

6a: Are the Results Positive or Concerning?
   Yes
   No

6b: Are the Results Positive or Concerning?
   Yes
   No

7a: 1. Provide Parental Education
     2. Schedule Extra Visit Within 1 Month
     3. Re-enter Algorithm at 1b

7b: 1. Schedule Next Preventive Visit
     2. Re-enter Algorithm at 1a

8: 1. Provide Parental Education
   2. Simultaneously Refer for:
      a. Comprehensive ASD Evaluation
      b. Early Intervention/Early Childhood Education Services
      c. Audiologic Evaluation
   3. Schedule Follow-Up Visit
   4. Re-enter Algorithm at 1b

Legend

= Start
= Action / Process
= Decision

FIGURE 1
Surveillance and screening algorithm: ASDs.
Surveillance and Screening Algorithm: Autism Spectrum Disorders (ASDs)

1a - Developmental concerns, including those about social skill deficits, should be included as one of several health topics addressed at each pediatric preventive care visit through the first 5 years of life. (Go to step 2)

1b – At the parents’ request, or when a concern is identified in a previous visit, a child may be scheduled for a “problem-targeted” clinic visit because of concerns about ASD. Parent concerns may be based on observed behaviors, social or language deficits, issues raised by other caregivers, or heightened anxiety produced by ASD coverage in the media. (Go to step 2)

2 – Developmental surveillance is a flexible, longitudinal, continuous, and cumulative process whereby health care professionals identify children who may have developmental problems. There are 5 components of developmental surveillance: eliciting and attending to the parents’ concerns about their child’s development, documenting and maintaining a developmental history, making accurate observations of the child, identifying the risk and protective factors, and maintaining an accurate record and documenting the process and findings. The concerns of parents, other caregivers, and pediatricians all should be included in determining whether surveillance suggests that the child may be at risk of an ASD. In addition, younger siblings of children with an ASD should also be considered at risk, because they are 10 times more likely to develop symptoms of an ASD than children without a sibling with an ASD. Scoring risk factors will help determine the next steps. (Go to step 3)

For more information on developmental surveillance, see “Identifying Infants and Young Children With Developmental Disorders in the Medical Home: An Algorithm for Developmental Surveillance and Screening” (Pediatrics 2006;118:409-420).

3a – 
- If the child’s age is <18 months, Go to step 5a
- If the child’s age is ≥18 months, Go to step 5b

3b – Is the Patient at Least 18 Months Old?

3c – For all children ages 18 or 24 months (regardless of risk factors), the pediatrician should use an ASD-specific screening tool. (Go to step 6b)

3d – If the child is not at least 18 months, Go to step 4

4 – In the absence of established risk factors and parental/provider concerns (score=0), a level-1 ASD-specific tool should be administered at the 18- and 24-month visits. (Go to step 5c) If this is not an 18- or 24-month visit, (Go to step 7b).Note: In the AAP policy, “Identifying Infants and Young Children With Developmental Disorders in the Medical Home: An Algorithm for Developmental Surveillance and Screening”, a general developmental screen is recommended at the 9-, 18-, and 24- or 30-month visits and an ASD screening is recommended at the 18-month visit. This clinical report also recommends an ASD screening at the 24-month visit to identify children who may regress after 18 months of age.

5a - If the child’s age is <18 months, the pediatrician should use a tool that specifically addresses the clinical characteristics of ASDs, such as those that target social-communication skills. (Go to step 6a)

5b – If the child’s age is ≥18 months, the pediatrician should use an ASD-specific screening tool. (Go to step 6a)

5c – Evaluate Social-Communication Skills

5d – Administer ASD-Specific Screening Tool

5e – For all children ages 18 or 24 months (regardless of risk factors), the pediatrician should use an ASD-specific screening tool. (Go to step 6b)

6a – When the result of the screening is negative, Go to step 7a

6b – When the result of the ASD screening (at 18- and 24-month visits) is negative, Go to step 7b

7a – If the child demonstrates risk but has a negative screening result, information about ASDs should be provided to parents. The pediatrician should schedule an extra visit within 1 month to address any residual ASD concerns or additional developmental/behavioral concerns after a negative screening result. The child will then re-enter the algorithm at 1b. A “wait-and-see” approach is discouraged. If the only risk factor is a sibling with an ASD, the pediatrician should maintain a higher index of suspicion and address ASD symptoms at each preventive care visit, but an early follow-up within 1 month is not necessary unless a parental concern subsequently arises.

7b – If this is not an 18- or 24-month visit, or when the result of the ASD screening is negative, the pediatrician can inform the parents and schedule the next routine preventive visit. The child will then re-enter the algorithm at 1a.

8 – If the screening result is positive for possible ASD in step 6a or 6b, the pediatrician should provide peer reviewed and/or consensus-developed ASD materials. Because a positive screening result does not determine a diagnosis of ASD, the child should be referred for a comprehensive ASD evaluation, to early intervention/early childhood education services (depending on child’s age), and an audiologic evaluation. A categorical diagnosis is not needed to access intervention services. These programs often provide evaluations and other services even before a medical evaluation is complete. A referral to intervention services or school also is indicated when other developmental/behavioral concerns exist, even though the ASD screening result is negative. The child should be scheduled for a follow-up visit and will then re-enter the algorithm at 1b. All communication between the referral sources and the pediatrician should be coordinated.


*Available at www.aap.org

FIGURE 1
Continued
Child Neurology Society practice parameter on screening and diagnosis of autism suggests that the following “red flags” are absolute indications for immediate evaluation:

- no babbling or pointing or other gesture by 12 months;
- no single words by 16 months;
- no 2-word spontaneous (not echolalic) phrases by 24 months; and
- loss of language or social skills at any age.

Pediatricians should become concerned (Step 2) if the answers to these questions reveal deficits or delays in milestones or if behaviors typical of ASD are observed during an office visit.

In older, more developmentally advanced children, including many with AS, surveillance questions (Step 2) may elicit concerns about social interaction difficulties related to more subtle communication problems, such as pragmatic language impairment and lack of understanding of nonliteral forms of communication (Figures of speech, humor, sarcasm, metaphor, etc), difficulty taking the perspective of another (resulting in inappropriate or offensive behavior, gullibility, and lack of common sense), and obsession with facts, details, or collections. Pragmatic language refers to the use of language in social interaction and includes instinctive rules governing factors, such as topic maintenance and turn taking in conversation, how sentences are made to fit in with the flow of a conversation, how unspoken premises are inferred, how degrees of formality and politeness are signaled, and prosody (modulation of the intonation, rhythm, volume, timing, and stress of the voice). The parents may note that the child lacks true friendships and is viewed as odd, eccentric, or “weird” by his ‘peers.

In addition, during the well-child visit, the PCP may try to interact with the patient by using a few simple strategies depending on the child’s age. For example, the PCP can note the response when calling the child’s name at the 12-month well-child visit, and/or the JA milestone of “following a point” can be elicited at the 12-, 18-, and 24-month well-child visits as part of routine developmental surveillance. In the latter, the pediatrician points to an object at a distance, such as a picture on the wall or a mobile, while making a verbal request for the child to look. Whereas a typically developing child would look in the direction of the point and then afterward engage in eye contact with the physician or the parent, a child with an ASD may appear to be oblivious to the PCP’s gesture and verbal request. This is true even if the PCP increases the intensity of the stimulus by calling louder, adding the child’s name, or touching the child’s shoulder first and then pointing and exclaiming, “Look!” The child may still fail to respond even if the parent repeats the maneuvers. With an older, higher-functioning child, the PCP may enter into conversation with the child to determine if he has difficulty interpreting a figure of speech, telling a joke, or explaining why a joke is funny. In addition, the PCP may ask a question or two about one of the child’s areas of interest to observe a response that is characteristic of AS, such as a long-winded, overly precise, or pedantic reply. Any of these responses should raise the concern of a PCP.

Each concern raised by a parent, other caregiver, or the pediatrician constitutes a separate risk factor, as does a positive family history of a sibling with an ASD (Step 2). To determine how to proceed, the pediatrician should assess the number of risk factors (Step 3). Possible scores include 0, 1, 2, 3, or 4.

1. If no concerns have been raised during the course of the preventive visit and the child is not the sibling of a child who has already been diagnosed with an ASD, then the PCP should proceed to Step 4. ASD-specific screening is indicated only if the visit is the 18- or 24-month preventive visit. See Step 5c below.

2. If the child’s only risk factor is having a sibling with an ASD, then the PCP should make sure the parents are aware of early signs of ASDs and continue to monitor carefully. If the parents call with a concern between scheduled routine preventive visits, the child should be seen within 1 or 2 weeks and reenter the algorithm at Step 1b for a “targeted visit” to address concerns about ASDs. If the score = 1 as a result of a single concern (parent, other caregiver, or PCP), the PCP should screen the child formally with a standardized tool; the choice of tool will depend on the child’s age (Step 3a) (see “Screening Tools for Implementation of Step 5”).

3. If 2 or more risk factors are identified, then the PCP should proceed directly to Step 8, which includes several activities that should be accomplished simultaneously and without delay.

### Screening for ASDs (Steps 5a–5c)

Physician estimates of the developmental status of children are much less accurate when only clinical impressions, rather than formal screening tools, are used, yet a minority of PCPs use formal developmental screening instruments, and few pediatricians specifically screen for ASDs. A standardized screening tool should be administered at any point when concerns about ASDs are raised spontaneously by a parent or as a result of clinician observations or surveillance questions about social, communicative, and play behaviors (Steps 5a and 5b). In the general developmental screening and surveillance policy statement discussed previously, the AAP also recommended administering a standardized autism-specific screening tool on all children at the 18-month preventive care visit (Step 5c). The AAP Autism Expert Panel responded to the statement with a com-
that suggested a repeat screening be performed at 24 months of age (Step 5c) to identify those who may regress after 18 months of age.

**Screening Tools for Implementation of Step 5**

A variety of general developmental screening tools are available to practitioners. General developmental screening tools are appropriate for use with unselected primary care populations and are likely to detect ASDs in many young children because of associated language and cognitive delays, but they do not differentiate children with ASDs from those with other developmental disorders, and data are not available on sensitivity for detection of ASDs. Tools to screen specifically for ASDs also have been designed (Table 3), but they have not yet been validated on children younger than 18 months. The PCP should remember that screening tools are likely to be overinclusive, so children with developmental and behavioral disorders other than ASDs also might have positive screening results. Similar to other developmental screening measures, ASD-specific screening tools may rely entirely on parent report, or they may require direct observation and engagement by the clinician. Parent-report tools often have the advantage of being brief, inexpensive, and practical in the office setting. The people who know the child best are surveyed and can describe the child’s behavior over time in a variety of settings rather than being constrained to sampling behavior in one setting at one point in time.

**Step 5a: Tools for Use in “at-Risk” Children Younger Than 18 Months**

Although several tools are in development for screening children younger than 18 months, none are available yet for routine clinical use. The Infant/Toddler Checklist from the Communication and Symbolic Behavior Scales Developmental Profile (which can be downloaded at www.brookespublishing.com/store/books/wetherby-csbudp/CSBSDP_Checklist.pdf) may be particularly well suited for identifying 6- to 24-month-old children who are at risk of ASDs, because it focuses on social and communication skills. It is anticipated that this and other screening tools under investigation as possible ASD-specific tools for use in infants younger than 18 months may prove valuable in identifying children at high risk and will become available to clinicians in the near future.

**Step 5b: Tools for at-Risk Children 18 Months and Older**

ASD-specific screening tools are available for children 18 months and older, and many of them are age specific. Recently, such tools have been classified as “level 1” or “level 2” screening tools. Level 1 screening tools are administered to all children within the context of a primary care medical home and are designed to differentiate children who are at risk of ASDs from the general population, especially those with typical development. Level 2 screening tools are used more often in early intervention programs or developmental clinics that serve children with a variety of developmental problems; they help to differentiate children who are at risk of ASDs from those at risk of other developmental disorders such as GDD or specific language impairment. Level 2 screening tools generally require more time and training to administer, score, and interpret than level 1 measures. There is considerable overlap between the concept of a level 2 screening tool and that of a diagnostic instrument. Level 2 screening measures may be used as part of a diagnostic evaluation, but they should not be used in isolation to make a diagnosis.

Properties of some level 1 and 2 ASD screening tools are reviewed in Table 3. Reported sensitivity and specificity values are included, but in most cases, sensitivity and specificity of the instruments have been determined only in clinical samples or in populations that included a mixture of clinical and population-based samples, and they must be interpreted with caution. Estimates of sensitivity and specificity of developmental screening tests may be unstable, and they are not the only criteria that should be used to assess validity. In low-prevalence conditions, such as ASDs, the positive predictive value of screening tools will be low even with good sensitivity and specificity, whereas the negative predictive value will be quite high. Many of the existing ASD-specific screening measures are being revised or further evaluated, and new tools are being developed to address some of their weaknesses.

Some measures, such as the Checklist for Autism in Toddlers (CHAT), Modified Checklist for Autism in Toddlers (M-CHAT), and Pervasive Developmental Disorders Screening Test-II Primary Care Screener, were designed specifically for early detection of ASDs in young children. The CHAT and M-CHAT are level 1 screening tools that are available at no cost to practitioners for use in primary care (Table 3).

For older children who are diagnosed later with AS, school personnel often raise concerns to the parents. Staff may then administer a published AS-specific tool. Although many level 2 screening tools have been marketed for use in older children who have been identified as being at risk of AS, further study is needed before any one of them can be recommended as superior to others. See Table 3 for characteristics of selected AS screening tools.

**Step 5c: Tools for Screening Children Without Risk Factors at the 18- and 24-Month Preventive Visit**

Level 1 ASD tools described in Step 5b also are appropriate for routine screening of young children without any identified risk.

Among the tools designed for screening the elementary school-aged population, only the Childhood Asperger Syndrome Test (CAST) has been assessed in a
<table>
<thead>
<tr>
<th>Screening Tool</th>
<th>Age</th>
<th>Format (No. of Items)</th>
<th>Time to Complete, min</th>
<th>Reported Sensitivity</th>
<th>Reported Specificity</th>
<th>Selected Key References</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1a CHAT</td>
<td>18–24 mo</td>
<td>Parent interview or questionnaire and interactive (parent: 9; clinician: 5)</td>
<td>5</td>
<td>0.18–0.38b; 0.65c</td>
<td>0.98–1.0b; 1.0c</td>
<td>Baron-Cohen et al.267, 272</td>
<td>Download: <a href="http://www.autismresearchcentre.com/tests/chat_test.asp">www.autismresearchcentre.com/tests/chat_test.asp</a></td>
</tr>
<tr>
<td>CHAT, Denver Modifications</td>
<td>18–24 mo</td>
<td>Parent interview or questionnaire and interactive (parent: 9; clinician: 5)</td>
<td>5</td>
<td>0.85c</td>
<td>1.0c</td>
<td>Scambler et al.273</td>
<td>CHAT scoring modifications; available in Scambler et al.273</td>
</tr>
<tr>
<td>Checklist for Autism in Toddlers-23 (CHAT-23)</td>
<td>16–86 mo (all had mental ages of 18–24 mo)</td>
<td>Parent interview or questionnaire and interactive (parent: 23, clinician: 5)</td>
<td>10</td>
<td>0.84–0.93c (part A); 0.74c (part B)</td>
<td>0.77–0.85c (part A); 0.91c (part B)</td>
<td>Wong et al.274</td>
<td>Combination of M-CHAT and CHAT items; protocol available in Wong et al.274</td>
</tr>
<tr>
<td>CAST</td>
<td>4–11 y</td>
<td>Questionnaire completed by parent (37)</td>
<td>10</td>
<td>0.88–1.0d</td>
<td>0.97–0.98d</td>
<td>Scott et al.275, Williams et al.276</td>
<td>Download: <a href="http://www.autismresearchcentre.com/tests/cast:.test.asp">www.autismresearchcentre.com/tests/cast:.test.asp</a></td>
</tr>
<tr>
<td>M-CHAT</td>
<td>16–48 mo</td>
<td>Questionnaire completed by parent (23)</td>
<td>5–10</td>
<td>0.85d</td>
<td>0.93d</td>
<td>Dumont-Mattheau and Fein,277, Williams et al.276</td>
<td>Download: <a href="http://www.dbpeds.org/media/mchat.pdf">www.dbpeds.org/media/mchat.pdf</a> or <a href="http://www.firstsigns.org/downloads/m-chat.pdf">www.firstsigns.org/downloads/m-chat.pdf</a>; for scoring: <a href="http://www.firstsigns.org/downloads/m-chat-scoring.PDF">www.firstsigns.org/downloads/m-chat-scoring.PDF</a></td>
</tr>
<tr>
<td>Pervasive Developmental Disorders Screening Test-II, Primary Care Screener (PDDST-II PCS)</td>
<td>18–48 mo</td>
<td>Questionnaire completed by parent (22)</td>
<td>10–15</td>
<td>0.92c</td>
<td>0.91c</td>
<td>Siegel260</td>
<td>Purchase: PsychCorp/Harcourt Assessment (<a href="http://www.harcourtassessment.com">www.harcourtassessment.com</a>)</td>
</tr>
<tr>
<td>Level 2 Asperger Syndrome Diagnostic Scale (ASDS)</td>
<td>5–18 y</td>
<td>Questionnaire completed by parent, teacher, or clinician (50)</td>
<td>10–15</td>
<td>0.85c</td>
<td>Myles et al.278, Campbell270</td>
<td>Purchase: Pro-Ed (<a href="http://www.proedinc.com">www.proedinc.com</a>)</td>
<td></td>
</tr>
<tr>
<td>Autism Behavior Checklist (ABC)</td>
<td>≥18 mo</td>
<td>Behavioral checklist completed by interviewer (57)</td>
<td>10–20</td>
<td>0.38–0.58c</td>
<td>0.35–0.97c</td>
<td>Krug et al.279</td>
<td>Purchase: Pro-Ed (<a href="http://www.proedinc.com">www.proedinc.com</a>) as part of the Autism Screening Instrument for Educational Planning (ASIEP-2)</td>
</tr>
<tr>
<td>Autism Quotient (AQ)–Adolescent Version</td>
<td>11–16 y</td>
<td>Questionnaire completed by parent (50)</td>
<td>15</td>
<td>0.89c</td>
<td>1.0c</td>
<td>Baron-Cohen et al.280</td>
<td>Download: <a href="http://www.autismresearchcentre.com/tests/aq_adolescent.asp">www.autismresearchcentre.com/tests/aq_adolescent.asp</a></td>
</tr>
<tr>
<td>Autism Spectrum Rating Questionnaire (ASQ)</td>
<td>6–17 y</td>
<td>Questionnaire completed by parent (27)</td>
<td>10</td>
<td>0.62–0.82c (parent); 0.65–0.70c (teacher)</td>
<td>Ehlers et al.281</td>
<td>Questions are included as an appendix in Ehlers et al.281</td>
<td></td>
</tr>
<tr>
<td>Childhood Autism Rating Scale (CARS)</td>
<td>&gt;2 y</td>
<td>Behavioral checklist completed by trained interviewer/observer (15)</td>
<td>Variable</td>
<td>0.92–0.98c; 0.94c</td>
<td>0.85c</td>
<td>Eaves and Milner282, Perry et al.283, Schopler et al.284, Sevin et al.285</td>
<td>Purchase: Western Psychological Services (<a href="http://www.wspublish.com">www.wspublish.com</a>)</td>
</tr>
<tr>
<td>Gilliam Asperger’s Disorder Scale (GADS)</td>
<td>3–22 y</td>
<td>Questionnaire completed by parent, teacher, or clinician (32)</td>
<td>10</td>
<td></td>
<td></td>
<td>Gilliam286, Campbell270</td>
<td>Purchase: Pro-Ed (<a href="http://www.proedinc.com">www.proedinc.com</a>)</td>
</tr>
</tbody>
</table>
TABLE 3  Continued

<table>
<thead>
<tr>
<th>Screening Tool</th>
<th>Age</th>
<th>Format (No. of Items)</th>
<th>Time to Complete, min</th>
<th>Reported Sensitivity</th>
<th>Reported Specificity</th>
<th>Selected Key References</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Krug Asperger’s Disorder Index (KADI)</td>
<td>6–21 y</td>
<td>Questionnaire completed by parent or clinician (32)</td>
<td>15–20</td>
<td>0.78&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.94&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Krug and Arick&lt;sup&gt;268&lt;/sup&gt;</td>
<td>Purchase: Pro-Ed (<a href="http://www.proedinc.com">www.proedinc.com</a>)</td>
</tr>
<tr>
<td>Pervasive Developmental Disorders Screening Test-II, Developmental Clinic Screener (PDDST-II, DCS)</td>
<td>18–48 mo</td>
<td>Questionnaire completed by parent (14)</td>
<td>10–15</td>
<td>0.73&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.49&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Campbell&lt;sup&gt;270&lt;/sup&gt;, Siegel&lt;sup&gt;269&lt;/sup&gt;</td>
<td>Purchase: PsychCorp/Harcourt Assessment (<a href="http://www.harcourtassessment.com">www.harcourtassessment.com</a>)</td>
</tr>
<tr>
<td>Pervasive Developmental Disorders Screening Test-II, Autism Clinic Severity Screener (PDDST-II, ACSC)</td>
<td>18–48 mo</td>
<td>Questionnaire completed by parent (12)</td>
<td>10–15</td>
<td>0.58&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.60&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Siegel&lt;sup&gt;269&lt;/sup&gt;</td>
<td>Purchase: PsychCorp/Harcourt Assessment (<a href="http://www.harcourtassessment.com">www.harcourtassessment.com</a>)</td>
</tr>
<tr>
<td>Screening Tool for Autism in Two-Year-Olds (STAT)</td>
<td>24–36 mo</td>
<td>Interactive, requires specific training (12)</td>
<td>20</td>
<td>0.92&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.85&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Stone et al&lt;sup&gt;289&lt;/sup&gt;, Stone et al&lt;sup&gt;290&lt;/sup&gt;</td>
<td>Author: Wendy Stone, PhD (<a href="mailto:triad@vanderbilt.edu">triad@vanderbilt.edu</a>)</td>
</tr>
<tr>
<td>Social Communication Questionnaire (SCQ) (formerly the Autism Screening Questionnaire [ASQ])</td>
<td>≥4 y</td>
<td>Questionnaire completed by parent (40)</td>
<td>5–10</td>
<td>0.85–0.96&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.67–0.80&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Berument et al&lt;sup&gt;291&lt;/sup&gt;, Rutter et al&lt;sup&gt;292&lt;/sup&gt;</td>
<td>Purchase: Western Psychological Services (<a href="http://www.wpspublish.com">www.wpspublish.com</a>)</td>
</tr>
</tbody>
</table>

The measures were selected on the basis of availability of some published psychometric properties (in English) with scoring instructions and pass/fail cutoffs or the equivalent.

<sup>a</sup> Level 1 tools are most likely to be used in primary care settings.

<sup>b</sup> Population-based sample.

<sup>c</sup> Clinical sample.

<sup>d</sup> Clinical and population-based samples.

large, unselected population as a level 1 screening tool. The authors concluded that the CAST is useful as a screening test for ASDs in epidemiologic research but that there is not enough evidence to recommend it for routine screening in the general population as part of a public health program. In addition, the AAP does not currently recommend universal screening of school-aged children with a level 1 AS-specific tool.

See Appendix 1 for reimbursement codes.

**Results of Screening (Steps 6a and 6b)**

If the screening result for an at-risk child is negative in Step 6a, the PCP should proceed to Step 7a, provide parent educational materials (such as the AAP brochure, “Is Your One-Year-Old Communicating With You?” or the AAP parent booklet, “Understanding Autism Spectrum Disorders”) and schedule an extra visit (Step 1b) within 1 month to address residual concerns. If the only risk factor is having a sibling with an ASD, an extra visit is not necessary unless the parents become concerned after the visit. When the screening result is negative for children without risk at the 18- or 24-month preventive visit (Step 6b), the PCP should proceed to Step 7b and schedule the next routine preventive care visit (Step 1a).

If the screening result is positive (Steps 6a or 6b) or 2 or more risk factors are present at Step 3, the PCP should proceed to Step 8, at which simultaneous activities should take place in an expedient manner. The PCP should consider the possibility that the child with a negative ASD screening result may have another developmental disorder that would warrant further investigation and referral to resources similar to those listed in Step 8.

When surveillance does not identify any risk factors and the visit is not an 18- or 24-month visit (Step 4), no screening is recommended, and the PCP may proceed directly to Step 7b.

**Step 8: Activities Needed When Multiple Risk Factors Are Present or When the ASD Screening Result Is Positive**

Activities described herein will depend on certain community characteristics, especially in regard to obtaining a comprehensive evaluation. Depending on the number of ASD experts in a given community, the interval wait for an appointment may be long. Thus, it is important that the PCP simultaneously accomplish all of Steps 8.1 through 8.4 while the family is waiting for a specialty appointment to confirm or rule out an ASD diagnosis.

**Step 8.1: Provide Parental Education**

If the PCP feels fairly certain that the child has a developmental disorder that falls somewhere in the autism spectrum, it will be helpful to give the parents reading materials. As discussed in the introduction to this report, the AAP has published “Understanding Autism Spectrum Disorders,” an educational booklet for parents with this intent. The comprehensive evaluation will progress more efficiently if the parents are more knowledgeable about the characteristic clinical symptoms of ASDs and can report them more accurately. Some PCPs are reluctant to share their concerns with parents, fearful that premature “labeling,” although it is tentative, might cause undue stress and anxiety on the part of the family. However, sincerity, honesty, and admitting uncertainty is appreciated by most parents. On the other hand, concealing a concern and taking a “wait-and-see” approach rarely is appreciated; in fact, this strategy often breeds parental discontent and, worse, resentment and anger. With the recent high visibility in the media, most parents (unlike before the 1990s) now are aware of ASDs and may suspect it and search the Internet for information. It is important that they receive peer-reviewed and consensus-driven information that is evidence based and that they understand how to interpret Web-site information that is not peer reviewed.

**Step 8.2.a: ASD Comprehensive Evaluation**

For some children, the diagnosis might be quite obvious to the PCP who is using the DSM-IV-TR criteria as a guide. In others, the diagnosis may be challenging, especially when externalizing behavioral symptoms are mild or variable and/or there are associated comorbid disorders. Ideally, the definitive diagnosis of an ASD should be made by a team of child specialists with expertise in ASDs. Unfortunately, teams are not available in every locale, and when they are, long waiting lists may exist. Most communities will have at least 1 pediatric subspecialist (eg, child neurologist, developmental pediatrician, psychiatrist) with at least some expertise in making an ASD diagnosis. Other professionals, such as child psychologists, SLPs, pediatric occupational therapists, and social workers with expertise in ASDs, can be helpful by performing independent evaluations, often using standardized tools that can assist in the diagnostic process, especially when no team or pediatric “expert” is available. Child psychologists with appropriate training and experience can make the diagnosis independently and often do so, especially in school systems. Recently, the American Speech-Language-Hearing Association published guidelines that stated that an SLP with expertise in ASDs can make the diagnosis independently when other resources are not available. Older children who first present with symptoms of AS after school entry often are first recognized and evaluated by the school district’s educational diagnostic team and subsequently, but unfortunately not always, referred to a health care professional.

If it seems fairly certain, on the basis of general developmental screening and/or available psychometric testing with standardized tools, that the child also has GDD or intellectual disability, then the PCP might order high-resolution karyotype and DNA testing for fragile X.
syndrome. If the child has clinical features (history, family history, physical examination) that are characteristic of a specific genetic or neurologic disorder that can be easily confirmed by a specific laboratory test, then the PCP may want to proceed with that test. On the other hand, the PCP may opt to refer the child to pediatric subspecialists for assistance with an etiologic workup and/or a search for coexisting conditions. Depending on availability and the nature of the concern(s), the PCP should consider a referral to a developmental pediatrician, a geneticist, and/or a child neurologist. See the next section for a more extensive discussion of the components of a comprehensive evaluation.

**Step 8.2.b: Early Intervention/Early Childhood Education Services**

As soon as an infant or toddler is suspected of having a delay or being at risk of a delay or developmental disorder such as an ASD, he should be referred immediately to an early intervention program (a government-subsidized public program designed to serve children with special needs and/or developmental delays from the time the problem is identified until the third birthday). If the child has had his third birthday, the referral should be made to the special education department in the local school. Among other professionals, assessment teams will almost always include SLPs and occupational therapists who can develop appropriate intervention plans without a categorical diagnosis. Intervention is important and often can be effective, even if it begins as generic speech therapy (ie, therapy that addresses most forms of language delay) and general developmental strategies. This intervention plan can be revised later to a more specific ASD intervention protocol (such as teaching JA) once the diagnosis is made. Experienced therapists often recognize ASD symptomatology and use strategies tailored to the child’s individual deficits, even without a definitive ASD diagnosis.

**Step 8.2.c: Audiology Evaluation**

All children with language delays, including those suspected of having ASDs, should undergo an audiology evaluation, even if the neonatal screening result was normal. This testing may be challenging to accomplish, because children with ASDs often are uncooperative for behavioral audiometry, the test most frequently used with toddlers. If the attempt is unsuccessful, an auditory brainstem response or brainstem auditory evoked-response test can be ordered; it is likely that sedation will be required. Sedation may be challenging, because some children with ASDs may respond paradoxically to sedatives.

**Steps 8.3 and 8.4: Schedule Follow-up Visit and Reenter Algorithm**

The child should be scheduled for a targeted follow-up visit within 1 month and reenter the algorithm at Step 1b to determine the status of the aforementioned referrals and to discuss any additional parental concerns once they have had the opportunity to read and learn more about ASDs.

**COMPREHENSIVE EVALUATION (SEE STEP 8.2.a)**

There are 3 major diagnostic challenges in the comprehensive assessment of a child with a suspected ASD: determining the child’s overall level of functioning; making the categorical diagnosis of an ASD; and determining the extent of the search for an associated etiology. To accomplish these 3 goals, a comprehensive evaluation should include the following components:

1. Health, developmental, and behavioral histories that include at least a 3-generation family pedigree and a review of systems.

2. Physical examination including a thorough search for dysmorphic features and neurologic abnormalities and a Wood’s lamp examination of the skin.

3. Developmental and/or psychometric evaluation (depending on age/skill level) to determine the child’s overall level of functioning and whether a discrepancy between motor-adaptive problem-solving and social communication skills is evident.

4. Determination of the presence of a categorical DSM-IV-TR diagnosis, preferably with standardized tools that operationalize the DSM criteria.

5. Assessment of the parents’ knowledge of ASDs, coping skills, and available resources and supports.

6. A laboratory investigation to search for a known etiology or coexisting condition guided by information obtained in Steps 1 through 5.

When appropriate, the evaluation should include information from multiple sources, because the child’s performance may vary among settings and caregivers. Depending on level of comfort, the PCP may opt to refer to an experienced pediatric subspecialist, such as a neurologist, geneticist, or developmental pediatrician, to further evaluate the child, especially when there is an abnormal neurologic finding, seizures, regression, dysmorphic features, and/or a complex family history.

Laboratory testing for children with ASDs (component 6 above) is controversial. Newer technology has been developed since publication of the 2001 AAP statement and technical report; however, some tests are not yet clinically available. Various specialists hold differing opinions about the definition of a “positive yield,” defined herein as a positive test result that indicates a known autism-related etiology (eg, a positive result on
DNA testing for fragile X syndrome or a karyotype revealing a mutation at 9q or 16p indicating tuberous sclerosis. They also promote varying clinical indications for extensive molecular testing and neuroimaging when the clinical validity of a positive finding is yet unknown in many cases. Some investigators have reported a positive yield when, in fact, the identified abnormality was nonspecific, did not relate to a known autism-related etiology, and did not affect counseling and/or management (eg, delayed myelinization on MRI). Medical symptoms should be evaluated on a case-by-case basis; rather than reflect an etiology, an abnormal test result may indicate that a child with an ASD has a coexisting condition (eg, a gastrointestinal disorder). Thus, an abnormal laboratory test result does not necessarily indicate a positive yield but may, indeed, indicate a condition that needs medical attention (see the AAP clinical report “Management of Children With Autism Spectrum Disorders”). Reporting it as a positive yield makes it difficult to translate research methodology into recommendations that will help the clinician in the care of any given patient.

The yield of an etiologic investigation may be more highly correlated with the presence or absence of coexisting GDD/MR (intellectual disability) rather than with an isolated ASD. In fact, the presence of autism in a cohort of children with GDD/MR (intellectual disability) decreased the chance of a positive yield. Depending on the population characteristics, specific test(s) studied, and the decision-making process by which they were ordered (ie, as a screening technique for all study patients with ASDs versus a targeted test indicated by a specific clinical finding), positive yields range from as low as 0% to as high as 25% to 40%, but most yield rates fall between 2% and 10%. It is difficult to compare studies because of variability in the workups, analysis in terms of GDD/MR or other phenotypic variables, and interpretation of positive test results (eg, delayed myelinization on MRI) or symptoms (eg, gastrointestinal) that are not definitively associated with ASDs.

Although the original ASD-specific consensus guidelines published between 1999 and 2001 have been helpful in guiding the etiology-search strategy in children with ASDs, the presence of coexisting GDD/MR (intellectual disability) in a cohort of children with ASD (especially severe GDD/MR or intellectual disability associated with dysmorphic features) is more highly correlated with a positive yield and a recognizable syndrome. Thus, guidelines that address the etiologic workup of children with GDD/MR also should be considered when evaluating a child with both an ASD and GDD/MR but not necessarily a child with an isolated ASD.

Among laboratory tests, high-resolution chromosome analysis by G-banding and molecular testing for fragile X syndrome have the highest yield in determining etiology in patients with ASDs. Some investigators have suggested a battery of additional screening cytogenetic and molecular studies for all patients with ASDs regardless of gender, presence or absence of coexisting GDD/MR, dysmorphic features, or family history. However, current data do not support extensive testing of all children with ASDs in clinical settings. Published studies have begun to address some of the newer molecular genetic techniques that have revolutionized genetic testing by detecting microdeletions, duplications, and rearrangements not visible with high-resolution chromosomal testing. Targeted FISH studies can be used to screen for deletions or duplications, such as those associated with chromosomes 15q and 22q. A relatively recent use of FISH technology is genome-wide subtelomere screening, which detects clinically significant abnormalities in 2.5% of individuals with unexplained GDD/MR. This technology can detect a wide variety of abnormalities, including some such as 22q13.3 deletion, that have been reported in a subset of children with ASDs. Several studies that examined the yield of subtelomere FISH screening in ASD failed to detect a single abnormality, which suggests that it may not be helpful in the routine evaluation of these patients. However, additional studies are needed. Comparative genomic hybridization-microarray analysis is a promising tool that may become standard of care in the future, but this technique has not been evaluated systematically in children with ASDs.

Screening neurologic tests also have been suggested—for example, electroencephalography (EEG [routine and/or prolonged sleep studies]) for all children with ASDs. Although nonspecific abnormalities have been found in most children, the significance of these abnormalities is not clear, and additional research is needed to determine if intervention is of any value. Thus, there is no evidence to support universal screening EEG without a clinical indication. An EEG should be considered for children who demonstrate clinical signs that might represent seizures and for children with clear language regression. However, EEGs in children that demonstrate “classic autistic regression” between 12 and 24 months are often nonspecific and not helpful in the diagnostic process. Previously published guidelines contain clear recommendations that screening MRIs on all children who present with ASDs, including those with isolated macrocephaly, are not necessary. Given the heterogeneity of ASDs, the likelihood of multiple etiologies, and the questionable clinical validity of an extensive battery of screening tests on all children with ASDs, more evidence is needed before a battery of genetic and neurologic testing becomes standard of care.

†Refs 20, 23, 33, 89, 97, 101, 105, 302, and 310.
‡Refs 20, 23, 33, 89, 101, 105, 302, and 310.
Although for the individual patient, it is important to differentiate an idiopathic ASD (with a recurrence rate of 5%–6% [range: 2%–8%]) from an ASD-associated syndrome that may have a higher or lower recurrence rate, there is no simple 1-size-fits-all search strategy. Instead, the search should be guided by clinical judgment based on history (eg, health, birth, developmental, behavioral, family) and clinical presentation (eg, comorbid MR, regression, seizures, neurodevelopmental findings, dysmorphic features, comorbid medical conditions). The importance of dysmorphic features and/or neurologic abnormalities in predicting a developmental finding, dysmorphic features, comorbid condition (eg, comorbid MR, regression, seizures, neurodevelopmental, behavioral, family) and clinical presentation (eg, comorbid MR, regression, seizures, neurodevelopmental findings, dysmorphic features, comorbid medical conditions). The importance of dysmorphic features and/or neurologic abnormalities in predicting a positive yield particularly has been emphasized. Familiar characteristics (eg. insurance status, concern about the child’s discomfort, or interest in pursuing a “no-stone-left-unturned” etiologic workup) also may affect parental decisions regarding the extent of the workup. Finally, the availability of technology, the need for and feasibility of sedation, managed care cost/benefit guidelines, and physician motivation each may play a role. There are certainly many advantages to having a diagnosis, including genetic counseling and provision of recurrence risks of known syndromes, the possibility of a specific treatment strategy, counseling regarding the natural history of a known disorder, anticipation of a later associated comorbid disorder, prevention of secondary disorders, availability of prenatal diagnosis, access to public support systems, access to syndrome-specific parent support groups, and, in some cases, the psychological benefits of knowing that empower parents to move on and focus on habilitative interventions.

A “search strategy” might be conceptualized as consisting of 3 levels.

1. Studies that should be considered for all young children with ASDs (ie, an audiology evaluation; however, school-based hearing screening may be adequate in the older child with AS and no significant language or learning deficits).

2. Studies that should be considered in all children with both an ASD and coexisting GDD/MR or intellectual disability (ie, high-resolution karyotype [650 bands] and DNA testing for fragile X syndrome). Although a high-resolution karyotype might reveal larger duplications, some clinicians believe that FISH testing for 15q duplications also might be indicated. In the future, a microarray analysis may replace high-resolution karyotyping. A methyl CpG-binding protein 2 (MECP2) analysis should be considered in females who present with regression and autistic features that are also consistent with Rett syndrome.

3. Targeted studies (eg, EEG, metabolic studies, MRI) should be considered when specific clinical findings are identified by history or physical examination (eg, seizures, cyclic vomiting and lethargy associated with mild illnesses and/or unusual odors, hypopigmented macules). Identification of more subtle indicators and their corresponding appropriate laboratory tests might be facilitated by referral to a geneticist, pediatric neurologist, and/or developmental pediatrician.

Ongoing multisite studies are investigating specific test protocols. Such evaluations are not recommended as clinical standard of care at this time until analysis of the data indicates which of the extended tests, if any, are indicated and for which ASD populations. These research protocols include many tests that are investigational, have unknown medical validity, and currently are not available for clinical use. Some of these tests include functional neuroimaging, immunologic studies, metabolic testing, fibroblast karyotypes, neureilin gene testing, mitochondrial gene sequencing, genomic microarrays, and identification of endophenotypes. Although these tests may not be relevant in clinical practice, they do have the potential to expand the fund of knowledge about ASDs, reveal more specific ASD subtypes, and provide a better understanding of coexisting disorders and future prognosis. As the fund of knowledge regarding genetic markers for ASDs expands and technology continues to become more sophisticated, the yield of these laboratory investigations may eventually prove to be useful in the routine clinical evaluation of children with idiopathic ASDs. For now, the existing dichotomy regarding the extent of testing in research versus clinical settings is challenging. Existing data do not support routine application of any particular test battery, nor do they suggest that tests currently under investigation be routinely performed on all children with ASDs at this time.

Prognosis

Although prognosis is one of the parents’ most pressing concerns at the time of diagnosis, it depends on many factors and usually cannot be predicted during early childhood, especially in children younger than 3 years. Important early predictors include JA skills, functional play skills, cognitive abilities, and severity of ASD symptoms. Recent studies have revealed that although most children diagnosed with AD retain their diagnosis at 9 years of age, many, especially those with PDD-NOS, improve, and a minority have optimal outcomes; that is, they have normal intelligence and function reasonably well in mainstream classrooms without an aid but still exhibit residual clinical signs of social awkwardness, restrictive interests, or mild, infrequent stereotypies. Some may show signs of ADHD, language-based learning disabilities, or other learning challenges. Poorer outcomes are associated with lack of JA by 4 years of age and lack of functional speech by 5 years of age. MR, seizures (especially with onset during adolescence), comorbid medical (eg, tuberous sclerosis) or psychiatric (eg, schizophrenia) disorders, and severe autistic symptoms, especially when associated with extreme “aloofness.” Factors associated with better out-
comes include early identification resulting in early enrollment in appropriate intervention programs and successful inclusion in regular educational and community settings with typically developing peers.

Adult outcomes seem to correlate better with level of cognitive-adaptive functioning than with the severity of autistic symptoms. People with normal intelligence/adaptive functioning and milder autistic symptoms generally have the best outcomes, those with MR or intellectual disability and severe autistic symptoms have the worst outcomes within the continuum, and those with normal cognitive-adaptive skills and severe autistic symptoms generally do better than those with MR or intellectual disability and mild autistic symptoms, reaffirming the contribution of intelligence rather than degree of atypicality (autistic symptoms). However, within the subgroup of children with normal intelligence, the degree of atypicality then becomes more important in determining prognosis. Many believe that people with AS have better outcomes than those with other ASDs. This may be true, because by definition, all those with AS have normal intelligence. One adult outcome study found that although those with AS tend to have a greater likelihood of earning a college degree than those with high-functioning autism/PDD-NOS, the college education did not significantly affect employment or marriage status.

Genetic Counseling

Genetic counseling regarding recurrence risk in siblings is important even when the etiologic evaluation is negative, because the recurrence risk is approximately 5% to 6% (range: 2%–8%) in a family with 1 child with an idiopathic ASD. The prevalence of abnormality in siblings is even higher, perhaps 20%, when the broader phenotype or milder constellation of similar social, communication, and behavioral abnormalities is considered. If there are already 2 siblings with ASDs in a family, it is likely that the recurrence risk for a strictly defined ASD in subsequent offspring is well above 8% and may approach 25%, but there is insufficient evidence to be more precise. It is important to discuss the recurrence risk promptly after diagnosis to provide parents with this information before they conceive another child. When an etiology is determined, the recurrence risk may be lower or higher than the risk in idiopathic ASD, depending on the syndrome or condition identified, and prenatal diagnosis may be possible.

GUIDANCE FOR PEDIATRICIANS REGARDING THE IDENTIFICATION AND EVALUATION OF CHILDREN WITH ASDs

In summary, most PCPs can expect to care for several children with ASDs in the context of the medical home. No two children with ASDs will be exactly alike; each will have his or her own constellation of diagnostic and management challenges. The PCP has an important role in the early identification of children with ASDs. PCPs should do the following:

1. Conduct surveillance at every well-child visit. Be a good listener and recognize the early subtle red flags that indicate the possibility of an ASD. Be especially vigilant for younger siblings of a child who has already been diagnosed with an ASD.

2. Screen at 18, 24, and 24–26 months and any other time when parents raise a concern about a possible ASD. Although no screening tool is perfect, choose and become comfortable with at least 1 tool for each age group and use it consistently. Before 18 months of age, screening tools that target social and communication skills may be helpful in systematically looking for early signs of ASDs.

3. If an ASD-specific screening result is negative but either the parents or the PCP remain somewhat concerned, then the PCP should schedule the child for an early, targeted clinic visit to address these persistent concerns.

4. Act on a positive screening result or when a child demonstrates 2 or more risk factors. Do not take a “wait-and-see” approach. Depending on the age of the child, simultaneously refer for all 3: comprehensive ASD evaluation; early intervention/early childhood education services; and an audiologic evaluation. Do not wait for a definitive diagnosis of an ASD to refer for developmental services; early intervention can be beneficial even if it targets the child’s unique deficits. The intervention strategy can be modified if needed when the child is determined to have an ASD.

The science of ASDs is expanding rapidly. Newer tools are under development and should become available to clinicians so that children can be screened and evaluated more efficiently and with greater accuracy in the future.

The reader is referred to the accompanying AAP clinical report, “Management of Children With Autism Spectrum Disorders,” to learn more about specific techniques and challenges in caring for children with ASDs within the context of a pediatric medical home.

APPENDIX 1: REIMBURSEMENT FOR SCREENING ACTIVITIES

Reimbursement for the administration of developmental and ASD-specific screening tools is an important aspect of screening. Developmental screening tests, including ASD-specific tests that are completed by a parent or nonphysician staff member and are reviewed and interpreted by the physician, can be billed appropriately by using Current Procedural Terminology (CPT) code 96110.

Tools that include a direct clinical observation component have the benefit of providing some potentially more objective information, and aspects of behavior that parents may not have noticed can be sampled. Extended
screening tests that include a direct testing component can be billed appropriately by using CPT code 96111.²⁴⁶

COUNCIL ON CHILDREN WITH DISABILITIES EXECUTIVE COMMITTEE, 2006–2007
Paul H. Lipkin, MD, Chairperson
J. Daniel Cartwright, MD
Larry W. Desch, MD
John C. Duby, MD
Ellen Roy Elias, MD
Chris Plauché Johnson, MD, MEd
Eric B. Levey, MD
Gregory S. Liptak, MD
Nancy A. Murphy, MD
Scott M. Myers, MD
Ann Henderson Tilton, MD

REFERENCES
27. Fombonne E, Zakarian R, Bennett A, Meng L, McLean-
42. Individuals With Disabilities Education Act Amendments. Pub L No. 105-17 (1997)
46. Laidler JR. US Department of education data on “autism” are not reliable for tracking autism prevalence. Pediatrics. 2005;116(1). Available at: www.pediatrics.org/cgi/content/full/116/1/e120


179. Redcay E, Courchesne E. When is the brain enlarged in autism? A meta-analysis of all brain size reports. Biol Psychiatry. 2005;58:1–9


199. Tuchman RF, Rapin I. Regression in pervasive developmental


257. Werner EE, Honzik MP, Smith RS. Prediction of intelligence and achievement at ten years from twenty months pediatric and psychologic evaluations. Child Dev. 1968;39:1063–1075


269. Siegel B. The Pervasive Developmental Disorders Screening Test II (PDDST-II). San Antonio, TX: Harcourt Assessment; 2004


318. Mantovani J. Regression and seizures. Presented at: the Spectrum of Developmental Disabilities XXVIII: Autism—From Kanner to Current; March 27–29, 2006; Baltimore, MD


RESOURCE FOR FAMILIES

Letters to the Editor

Beyond Munchausen Syndrome by Proxy

To the Editor.—

As experts in Munchausen syndrome by proxy (MSBP) maltreatment, we are writing in response to “Beyond Munchausen Syndrome by Proxy: Identification and Treatment of Child Abuse in a Medical Setting.” We commend the American Academy of Pediatrics for addressing this form of maltreatment, which is underrecognized. However, we do have several concerns about this position statement. Of these concerns, the most important is the remark that, although multidisciplinary input is important, the physician is the only professional who can actually make the diagnosis of MSBP.

We differ with this conclusion for several reasons. Most basically, MSBP is a much wider phenomenon than just “a form of child abuse taking place in a medical setting.” Manifestations of MSBP can be seen in schools, mental health facilities, nonprofit organizations, churches, the legal system, child protection agencies, the home, and the community at large. Likewise, physical symptoms are only a part of the spectrum of MSBP, with other kinds of “problems” (eg, psychological symptoms and other behaviors) that are exaggerated, fabricated, or induced in some cases.

The methodology involved in confirming or disconfirming MSBP includes the gathering of records from many sources (not just health care), conducting specialized interviews of a variety of professionals and nonprofessionals, and performing an overall information analysis, the last conducted by, or with the assistance of, a professional who has specialized knowledge and experience in this field. The task of identifying the discrepancies that may be present among all these sources, and answering the medical questions that arise, often requires input and clarification from physicians, but the investigation as a whole need not be done solely by them. In fact, most physicians do not have the time required for a thorough, appropriate investigation. Although the authors stated that “child abuse is a pediatric diagnosis,” it is not solely a pediatric medical diagnosis. Issues of motivation and circumstance are always important in determining that a given physical condition is the result of deliberate abuse (or neglect), as opposed to accident, ignorance, or other possible causes. All states and many other nations reflect this fact by assigning maltreatment investigations to child protective agencies and, within hospitals, to multidisciplinary/multiagency child protective teams.

The statement that “psychologists, social workers, and others are not in a position to make or confirm this diagnosis” overlooks the legal requirement for these professionals, and others, to report suspicions of child maltreatment to child protective authorities. These professionals receive training in the assessment of child maltreatment, may incur legal penalties if they fail to report, and must act in accord with their professional ethics. Given the overall shortage of professionals with expertise in MSBP, it is often necessary that a counselor, social worker, or other nonphysician make the assessment and lead the multidisciplinary/multiagency team. In such circumstances, involved professionals should make every effort to work with an MSBP professional of whatever discipline and/or to obtain education from an MSBP professional in this specialized kind of maltreatment.

We believe that it can be dangerous to delay the reporting of suspicions of child maltreatment, including MSBP, as suggested by the authors’ “list of possible interventions.” Our experience has been that many situations of MSBP respond very poorly to “individual and/or family therapy,” which may even be contraindicated as an early step in MSBP.Attempts by a physician to “gatekeep” or “monitor” are thwarted easily by perpetrators of MSBP who involve multiple caregivers (“doctor-shop”) or even flee to another area when suspected. The extent of MSBP maltreatment perceived by the primary caregiver may be only the tip of the iceberg.
of the iceberg if, unknown to the pediatrician, other providers or agencies are also being deceived. Absent the supervision and power of the courts, it is likely that perpetrators will not tell the whole truth, genuinely cooperate with medical gatekeeping, or even remain available to the physician.

Marc D. Feldman, MD
Department of Psychiatry and Behavioral Medicine
University of Alabama
Tuscaloosa, AL 35487

Michael J. Light, MD
Department of Pediatrics
Miller School of Medicine
University of Miami
Miami, FL 33136

Louisa J. Lasher, MA
Monticello, GA 31064

Mary S. Sheridan, PhD, ACSW
Department of Social Work
Hawaii Pacific University
Honolulu, HI 96813

REFERENCES

In Reply.—

In committing to write a report on the falsification of pediatric medical conditions (Munchausen syndrome by proxy [MSBP]), the Committee on Child Abuse and Neglect knew that it was wading into a definitional swamp of confusing notions about this type of child abuse. The response from Feldman, Light, Lasher, and Sheridan, all of whom are respected and competent professionals, sheds light on the necessity for ongoing debate.

The committee set out with 2 objectives. One was to clarify the sometimes bewildering terminology that is applied to these cases. We pointed out that the issue for the medical provider is the falsification of a pediatric condition, more than the state of mind or motivation of the perpetrator. We are concerned about the harm that can be done to a child in the medical setting with medical personnel as the unwitting instrument of the abuse. Although some authors have applied the term to someone who lies about the nonexistent illnesses of a fictitious child over the Internet or a woman who falsifies the illness of her dog, our discussion was limited to medical presentations, as in the original patients with Munchausen syndrome. In such cases, the medical professional is, in fact, the only member of the team who can determine if the condition truly exists and thus “make” the diagnosis. Nowhere is it contended that the medical professional is able to investigate the case without help, nor is it felt that only the physician must lead the team. We certainly agree with the authors that child abuse is not only a medical diagnosis. Nonetheless, we maintain that the diagnosis of factitious disease remains a medical one.

Our other objective was to remind pediatric providers of their role in the ongoing management of these cases. Many presentations that suggest MSBP fall short of a report to child protective services, and many of those reported need not be removed from the home. One response will not be appropriate for all. To this end, we offered a number of progressively restrictive management options, with multidisciplinary team involvement high on the list. Clearly, at the more severe end of the spectrum, this type of abuse is as serious as any and use of the child protection system is absolutely necessary.

In their response to our article, Feldman et al again underline the importance of a multidisciplinary approach to the difficult diagnosis of MSBP and remind readers that there is an “overall shortage of professionals with expertise in MSBP.” We can certainly not disagree with either statement. What was stressed in our article is the need for pediatricians to remember that parents’ histories may be unreliable and to maintain vigilance.

John Stirling, Jr, MD, FAAP
Carole Jenny, MD, MBA, FAAP
For the American Academy of Pediatrics Committee on Child Abuse and Neglect

REFERENCES

Unconventional Solution: A Marketing Approach to Solving Health Literacy Disparities

To the Editor.—

We read with interest the recent report from Sanders et al in regard to the association between caregiver’s health literacy and the use of child health services. Their finding raises an important question about the true role that health literacy plays in health outcomes.

A large body of research suggests that health illiteracy does not completely account for suboptimal outcomes. For example, some studies found that illiteracy...
was associated with reduced knowledge about medications,\(^2\) participation in research,\(^3\) and compliance with medical regimens,\(^4\) which indicates that illiteracy might be an important factor in poor disease management. However, other researchers have failed to identify an association between illiteracy and reduced odds of adherence to prescribed medication regimens\(^5\)–\(^7\) or poor refill adherence.\(^8\) In addition, limited health literacy has been associated with poor knowledge of a medication but not with self-reported adherence to taking that medication.\(^2\)

Undoubtedly, literacy must play an important role in compliance, because it provides a framework for action. Information is understood by the individual; consequently, there is a possible decision on which to act on the basis of this information. However, the available evidence indicates that low compliance occurs among both literate and illiterate individuals and among people with and without health insurance coverage. The “critical component” that triggers the necessary actions for good compliance to disease management, discharge instructions, or behavioral treatments is missing. As an example, we might consider the empirical experiences of people who continue to smoke despite the fact that they know, very well, the consequences of this unhealthy behavior. Another illustration of the problem includes people who continue unhealthy eating habits and sedentary lifestyles although they are fully aware that they should modify their diet and increase exercise to reduce their risk for diabetes and other cardiovascular problems. These experiences suggest that even when the information is delivered and understood, the “right message” needed to motivate positive actions and increase compliance is “communicated” but, unfortunately, “not received.”

To find the blueprint for transmission of the right message, we might look no further than the expertise of the professionals in advertising and marketing. Although marketing is truly an unconventional approach, we must take notice of the impressive ability of marketing experts to motivate both literate and illiterate persons to take action on the basis of advertising messages. These experts on consumer behavior are capable of formulating messages that can motivate (“trigger”) individuals to purchase items and services that are either not needed or not perceived as needed. Marketing is a new discipline with <100 years’ existence. It has drawn its strengths from other disciplines such as communication, sociology, anthropology, and psychology. The process of motivating consumers to purchase a service or product is simply explained by looking at one of the most quoted formulas in advertising and marketing since the 1950s, known as the AIDA technique. “AIDA” is an acronym for attention, interest, desire, and action. AIDA or newer variations are applied to virtually every marketing message to motivate consumers to action, usually reflected in the purchase of a product or service. This unconventional solution that includes adopting a “marketing approach” to communicate with low-literacy populations might be the missing link needed to increase compliance or adherence with disease management, discharge instructions, or instructions for behavior modification. Without discarding the role that health literacy might play in health outcomes, perhaps its influence is not as we imagine or health literacy measurement needs to be determined by using new and more precise venues.

**REFERENCES**


**In Reply.—**

We agree with the thoughtful comments from Leiner and Handal regarding the role of marketing science as a “missing link” in addressing literacy-related disparities in health outcomes. Indeed, the Institute of Medicine’s 2004 report on health literacy called for greater attention to the impact of both commercial and social marketing on health behaviors,\(^1\) and recent studies have noted the significant role that advertisers play in shaping the health behaviors of very young children.\(^2\) In fact, social learning theory suggests that marketing strategies are often most successful when they
pay less attention to consumer understanding and more attention to consumer motivation, self-efficacy, and behavior. This may help explain why, in our study of an underserved population, parent health literacy (a proxy for understanding health information) did not correlate strongly with child health care utilization (one type of health behavior).

Nonetheless, fully informed participation in health decisions remains a central tenet of the US health care system. Many aspects of routine pediatric care (including newborn screening, medication administration, and vaccination) require that children and their family members have a full understanding of risks, benefits, and options. Attending to the health literacy needs of the population requires an interdisciplinary approach that includes cognitive psychology, anthropology, sociology, and marketing. Health literacy researchers interested in health behavior change and a more accessible health care system should actively engage our colleagues from across the campus, including the business school. Written medication instructions, disease management plans, and informed-consent documents could all benefit from a redesign at the hands of the same marketing scientists who set in motion the logos, tag lines, and jingles that drive our purchasing behaviors.

Lee M. Sanders, MD, MPH
James D. Wilkinson, MD, PhD
Valerie Thompson, MD
Department of Pediatrics
Miller School of Medicine
University of Miami
Miami, FL 33136

REFERENCES
doi:10.1542/peds.2007-2502

Parental Attitude to Lumbar Puncture for Children With Fever and Seizure

To the Editor.—

I read with interest the case presented by Golnik of a 1-year-old girl with pneumococcal meningitis who presented with fever and a febrile seizure. It can be inferred, although it was not mentioned explicitly, that the child in this case report had not received antibiotics for any other reason during the days preceding her presenting illness.

In the presence of a negative blood-culture result, lumbar puncture proved to be the crucial diagnostic test. Parental opposition, despite a discussion of the risks and benefits of the procedure, caused a delay in the initiation of antibiotic treatment, which could have resulted in more serious long-term consequences. The way physicians introduce lumbar puncture to parents, not as a mere option but as a crucial diagnostic procedure in this and similar instances, could be an important determinant of their acceptance. While mentioning the risks of the procedure itself (which should, in experienced hands, be exceedingly low), perhaps more importantly, the risks of not performing it in such a situation have to be stressed. How can we do better when it comes to convincing parents of the necessity of performing such a procedure in a timely fashion, and what should we do if they remain opposed? Should we err on the side of caution and start treatment for presumed meningitis, or is watchful waiting, as done in this case, the only option?

This child had a neutrophil leukocytosis with an elevated percentage of bands on her initial blood count. In retrospect, although the initial clinical picture was in favor of a viral illness, such neutrophil leukocytosis, particularly in the presence of a left shift, might have been one indicator of a bacterial illness. Could such leukocytosis with a left shift, even in the absence of atypical seizure or any other clinical signs of meningitis, have added to the weight of evidence in support of performing a lumbar puncture, tipping the benefit/risk equation in favor of a lumbar puncture from a parent’s perspective?

Ayman Habiba, MD
Consultant Pediatrician
Al Mowasat Hospital
Salmiya 22077, Kuwait

REFERENCES
1. Golnik A. Pneumococcal meningitis presenting with a simple febrile seizure and negative blood culture result. Pediatrics. 2007; 120(2). Available at: www.pediatrics.org/cgi/content/full/120/2/e428
doi:10.1542/peds.2007-2368
In Reply.—

I am grateful for the letter from Dr Habiba regarding parental acceptance of lumbar puncture (LP) and bandemia as a predictor of meningitis. The child in the case described had not received antibiotics before evaluation. Parental opposition to LP was one of the many factors that influenced the procedure decision. In addition, the clinical evaluation and debate surrounding the indications to perform LP in children aged 12 to 18 months with simple febrile seizure contributed to the decision.1–5

Medical decision-making is increasingly shared by the physician and patient as equal partners.6 The American Academy of Pediatrics Committee on Bioethics has suggested that (when obtaining informed permission) physicians provide information in understandable language, consider the patient/parent’s understanding and capacity to make necessary decisions, and assure the family that they have the freedom to choose between medical alternatives.6 If parents refuse a necessary procedure, physicians should focus on providing appropriate medical care and be prepared to seek legal intervention if the refusal places the child at substantial medical risk.6 Specifically, if parents refuse permission for LP, the American Academy of Pediatrics Committee on Bioethics suggested obtaining parental permission to initiate treatment on the basis of reasonable clinical judgment rather than postponing care or risking liability for performing the procedure without parental consent.6

Regarding the bandemia with neutrophil leukocytosis, bands indicate the rapid production of white blood cells and may be a predictor of bacterial meningitis.7 Bandemia with neutrophil leukocytosis should provoke strong consideration of a bacterial etiology (urinary tract infection, bacteremia, meningitis, etc.). Such evidence of a bacterial etiology can be communicated to parents when discussing the risks and benefits of LP.

Allison Golnik, MD
Division of General Pediatrics
Department of Pediatrics
University of Minnesota
Minneapolis, MN 55455

REFERENCES

doi:10.1542/peds.2007-2536

Possible Sequelae of Sustained Lung Inflation in Resuscitation of Preterm Infants

To the Editor.—

I read with interest the recent report by te Pas and Walther,1 which showed a decrease in intubation and bronchopulmonary dysplasia in preterm infants of >28 weeks’ gestation by using sustained inflation and early nasal continuous positive airway pressure.

The first breaths of life in normal term infants are characterized by prolonged expiratory phases, associated with high positive intrathoracic pressure, interspersed with brief inspiratory components.2 Vyas et al3 demonstrated that sustained lung inflation during the resuscitation of 10 asphyxiated infants produced a large increase in the tidal volume and the functional residual capacity. However, a long inspiratory time is associated with a significant increase in air leak, as demonstrated in a Cochrane review that included 5 studies, with 694 infants recruited.4 In addition, mortality before hospital discharge was increased and reached borderline statistical significance. These studies were conducted before the introduction of antenatal steroids, postnatal surfactant, and the use of synchronized modes of ventilatory support.4 However, none of the studies of preterm infants examining sustained inflations during resuscitation1,5,6 measured respiratory parameters such as tidal volume, inspiratory time, or inflating pressure, and small numbers may prevent statistical significance in either benefits or complications. Harling et al1 illustrated the changes in PCO2, Po2, mean pH, median fraction of inspired oxygen, and median peak inspiratory pressure over the first 24 hours after either conventional or sustained lung inflation. Were these parameters available to the authors?

Harling et al showed no improvement in outcome after sustained inflations (5 seconds) and suggested that developmentally immature lungs that are deficient in surfactant may be unable to respond to this inflation maneuver. te Pas and Walther1 showed an increase in complications such as pneumothorax and severe intraventricular hemorrhage in the infants in the intervention group, although they did not reach statistical significance. I wondered whether the authors could provide additional details on the outcome of the subgroup of infants who were <28 weeks’ gestation, including com-
Eleanor J. Molloy, MB, PhD, FRCPI
Consultant Neonatologist
National Maternity Hospital
Dublin 2, Ireland

REFERENCES

doi:10.1542/peds.2007-2497

In Reply.—

We thank Dr Molloy for her interest in our randomized, controlled trial (RCT).1 She makes several comments about prolonged inflations, but her main question is unclear.

She refers to a Cochrane review2 and writes, “However, a long inspiratory time is associated with a significant increase in air leak.” This review analyzed trials of ventilation that used long inspiratory times for infants with respiratory distress syndrome in the NICU. It was not about prolonged inflation during delivery room resuscitation and, therefore, is not appropriate in the context of this trial.1,2

We are not sure why Molloy refers to a retrospective study by Lindner et al1 and our RCT and states that “small numbers may prevent statistical significance in either benefits or complications.” Both studies were large enough to show significant benefit with a decreased incidence of intubations.

Molloy refers to a study by Harling et al,4 which showed no difference in outcome between an initial inflation of 5 versus 2 seconds at the start of resuscitation for preterm infants, but as Molloy noted correctly, that sample size was small.

We did not record changes in blood gasses, median fraction of inspired oxygen (FIO2), or median peak pressures for the first 24 hours. Nevertheless, we did report the first blood gas in the NICU and the maximum FIO2 with no difference between the 2 groups.1 There was also no difference in the maximal peak pressure of ventilated infants (20.1 ± 2.4 [early functional respiratory capacity intervention (EFURCI) group] vs 19.2 ± 4.3 [conventional group]) cm H2O; P = .2.

We think Molloy misread the incidence of pneumothorax. The results (page 325, paragraph 2, lines 18–20) showed that there were fewer pneumothoraces in the EFURCI group (1 in 104 [EFURCI] vs 7 in 103 [conventional]; P = .069).1 There was a nonsignificant increase in severe intraventricular hemorrhage (IVH) in the EFURCI group (7 in 104 [7%] [EFURCI] vs 3 in 103 [3%] [conventional]; P = .3).1 The incidence of severe IVH in our NICU before the trial (year 2003–2004) was 7%. In addition, in our RCT the incidence of all IVHs was significantly higher in the conventional group (14 in 104 [14%] [EFURCI] vs 28 in 103 [27%] [conventional]; P = .016). Interestingly, Lindner et al1,5 showed no differences in IVH in either their retrospective study or RCT.

As requested, in Table 1 we show outcomes for infants <28 weeks’ gestation. For comparison, we also show the incidence of complications in the 2 years before the trial. As readers will appreciate, this trial was not designed or powered to investigate this subgroup.

Arjan B. te Pas, MD
Frans J. Walther, MD, PhD
Division of Neonatology
Department of Pediatrics
Leiden University Medical Center
2300 RC Leiden, Netherlands

<table>
<thead>
<tr>
<th>Secondary Outcome</th>
<th>EFURCI Group (n = 20)</th>
<th>Conventional Group (n = 19)</th>
<th>Univariate Analysis, P (Fisher’s exact test)</th>
<th>Before Trial, 2003–2004</th>
<th>Conventional Group (n = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumothorax, n (%)</td>
<td>0 (0)</td>
<td>1 (5)</td>
<td>.5</td>
<td>4 (7)</td>
<td></td>
</tr>
<tr>
<td>IVH grades 3 and 4, n (%)</td>
<td>4 (20)</td>
<td>1 (5)</td>
<td>.3</td>
<td>10 (19)</td>
<td></td>
</tr>
<tr>
<td>Cystic PVL, n (%)</td>
<td>1 (6)</td>
<td>2 (11)</td>
<td>.99</td>
<td>1 (2)</td>
<td></td>
</tr>
<tr>
<td>ROP above grade 3, n (%)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>—</td>
<td>1 (3)</td>
<td></td>
</tr>
<tr>
<td>NEC at least stage 2, n (%)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>—</td>
<td>2 (4)</td>
<td></td>
</tr>
<tr>
<td>Mortality, n (%)</td>
<td>2 (10)</td>
<td>2 (11)</td>
<td>.99</td>
<td>18 (33)</td>
<td></td>
</tr>
<tr>
<td>BPD/moderate-severe, n (%)</td>
<td>5 (28)</td>
<td>7 (41)</td>
<td>.5</td>
<td>23 (64)</td>
<td></td>
</tr>
</tbody>
</table>

PVL indicates periventricular leukomalacia; ROP, retinopathy of prematurity; NEC, necrotizing enterocolitis; BPD, bronchopulmonary dysplasia.

*Percentage of survivors.
The recent Moya et al article1 reported on a meta-analysis as opposed to patient-level analysis despite clear access to the databases as evidenced by the publication authorships.

The results were misleading for the following reasons.

1. There were clear differences between the Safety and Effectiveness of Lucinactant Versus Exosurf in a Clinical Trial of Respiratory Distress Syndrome in Premature Infants (SELECT)2 and Surfaxin Therapy Against Respiratory Distress Syndrome (STAR)3 populations with respect to birth weight (36 g mean lower in the STAR study), gestational age (1 week mean lower in the STAR study), prenatal steroid use (8% higher in the STAR study), and Apgar score (1 unit higher in the STAR study). This is best addressed by performing a combined analysis (at the neonate level) to control for these factors.

2. The different randomization allocations were a weak reason to not pool the data. The use of a meta-analysis is not efficient when all such baseline data are available and there are study and treatment imbalances.

3. The authors did not state if 1-year adjusted survival was the primary efficacy end point or how the analysis was to be performed. A logistic regression model is advised for a 1-year survival end point to diffuse the impact of statistical test choice (log rank versus Wilcoxon-Gehan, which are well known to handle death timing in different manners).

4. The life tables shown in their Fig 2 are quite different for the 2 studies. The Fig 2 labels are misleading relative to the text and the contents; Fig 2A looks like the pooled life table without colfosceril. The authors did not cite or explain the clear survival differences between the studies.

5. The survival rate with poractant seemed more favorable than that with the other 2 animal-based surfactants, but the authors still tried to make a comparison of combined animal-based surfactants versus lucinactant without any justification. They incorrectly assumed that all animal-based surfactants can be pooled; they also did not examine lucinactant variability across the 2 studies. Thus, this conclusion was without appropriate methodologic basis.

6. Numerically, the 1-year survival rate with poractant (21.9% without imputation) is, in fact, the numeric best of the 4 therapies (22.6% averaging nonimputed lucinactant across studies); a patient-level multivariate analysis is recommended to adjust for the study and treatment-group imbalances.

7. The STAR authors’ claim that the 1-year “corrected” lucinactant was superior to poractant (P = .04) is, at best, a posthoc comparison, because the retrospective power is only 29% to rule out a 5% disadvantage even using the 651 neonates on lucinactant versus the 128 neonates on poractant.

8. The STAR study itself only had 23% power to rule out a 5% disadvantage before the study started (unlikely that the study was planned to detect this disadvantage); however, it did have 80% power to rule out a 9.3% retrospective disadvantage but was marginally significant to rule out a 5% disadvantage.

The publication, therefore, was misleading. The claims were exaggerated beyond sound statistical analysis principles.

Philip T. Lavin, PhD
Averion International Corp
Southborough, MA 01772

Financial Disclosure: Dr Lavin acts as a consultant for Chiesi Farmaceutici SpA.

REFERENCES

Meta-Analysis Combining 2 Previously Reported Trials on Respiratory Distress Syndrome in Neonates
To the Editor.—

The recent Moya et al article1 reported on a meta-analysis as opposed to patient-level analysis despite clear access to the databases as evidenced by the publication authorships.

The results were misleading for the following reasons.

1. There were clear differences between the Safety and Effectiveness of Lucinactant Versus Exosurf in a Clinical Trial of Respiratory Distress Syndrome in Premature Infants (SELECT)2 and Surfaxin Therapy Against Respiratory Distress Syndrome (STAR)3 populations with respect to birth weight (36 g mean lower in the STAR study), gestational age (1 week mean lower in the STAR study), prenatal steroid use (8% higher in the STAR study), and Apgar score (1 unit higher in the STAR study). This is best addressed by performing a combined analysis (at the neonate level) to control for these factors.

2. The different randomization allocations were a weak reason to not pool the data. The use of a meta-analysis is not efficient when all such baseline data are available and there are study and treatment imbalances.

3. The authors did not state if 1-year adjusted survival was the primary efficacy end point or how the analysis was to be performed. A logistic regression model is advised for a 1-year survival end point to diffuse the impact of statistical test choice (log rank versus Wilcoxon-Gehan, which are well known to handle death timing in different manners).

4. The life tables shown in their Fig 2 are quite different for the 2 studies. The Fig 2 labels are misleading relative to the text and the contents; Fig 2A looks like the pooled life table without colfosceril. The authors did not cite or explain the clear survival differences between the studies.

5. The survival rate with poractant seemed more favorable than that with the other 2 animal-based surfactants, but the authors still tried to make a comparison of combined animal-based surfactants versus lucinactant without any justification. They incorrectly assumed that all animal-based surfactants can be pooled; they also did not examine lucinactant variability across the 2 studies. Thus, this conclusion was without appropriate methodologic basis.

6. Numerically, the 1-year survival rate with poractant (21.9% without imputation) is, in fact, the numeric best of the 4 therapies (22.6% averaging nonimputed lucinactant across studies); a patient-level multivariate analysis is recommended to adjust for the study and treatment-group imbalances.

7. The STAR authors’ claim that the 1-year “corrected” lucinactant was superior to poractant (P = .04) is, at best, a posthoc comparison, because the retrospective power is only 29% to rule out a 5% disadvantage even using the 651 neonates on lucinactant versus the 128 neonates on poractant.

8. The STAR study itself only had 23% power to rule out a 5% disadvantage before the study started (unlikely that the study was planned to detect this disadvantage); however, it did have 80% power to rule out a 9.3% retrospective disadvantage but was marginally significant to rule out a 5% disadvantage.

The publication, therefore, was misleading. The claims were exaggerated beyond sound statistical analysis principles.

Philip T. Lavin, PhD
Averion International Corp
Southborough, MA 01772

Financial Disclosure: Dr Lavin acts as a consultant for Chiesi Farmaceutici SpA.

REFERENCES


doi:10.1542/peds.2007-2375

To the Editor—

In the recent article “One-Year Follow-up of Very Premature Infants Who Received Lucinactant for Prevention of Respiratory Distress Syndrome: Results From 2 Multicenter Randomized, Controlled Trials,”1 Moya et al re-examined 2 different trials (Surfaxin Therapy Against Respiratory Distress Syndrome [STAR]2 and Safety and Effectiveness of Lucinactant Versus Exosurf in a Clinical Trial [SELECT]3) on respiratory distress syndrome and performed a meta-analysis by pooling the 2 studies together.

Although meta-analysis can be a valuable way of summarizing evidence, it suffers from the disadvantage, compared with the prespecified analysis of individual randomized clinical trials, of being retrospective. This is well understood in a regulatory context. For example, the Committee for Proprietary Medicinal Products guidelines state:

To minimize the opportunity for a retrospectively specified meta-analysis protocol . . . to be data dependent, the primary specifications and definitions set up in the individual studies should be followed. The credibility of the meta-analysis will depend on the degree of adherence to the specifications in the individual study protocols.4

Of course, standards for publication are not as rigorous as for regulatory submission. Nevertheless, when authors exchange a prespecified analysis with one determined after seeing the data, it behooves them to choose wisely and justify their choice scientifically.

We believe that this meta-analysis failed on these counts.

First, from a clinical point of view, it is irrelevant that a premature infant survives a few more weeks in a treatment group or in another, because the key point, rather, is whether the infant survives the critical period of vulnerability.

To the extent that some reasonable follow-up time can be agreed upon by which, if a child is alive, he or she is highly likely to survive to adulthood, a binary outcome (alive/dead) at this time point becomes an appropriate end point on which to focus. Therefore, an appropriate way of analyzing such data is to perform a logistic regression, not a survival analysis.

Typically, survival analysis is used when faced with data for which, by a reasonable time horizon, most subjects would die, and longer survival, therefore, is important. In such cases the focus is on time to event rather than the probability of the event itself occurring. Under such circumstances, the use of a survival analysis will permit recovery of relevant information that otherwise would be lost.

The same authors in the original publication (ie, STAR study2), followed a strategy consistent with this argument and used the binary outcome of alive/dead at 28 days as primary clinical outcome to assess treatment differences. Also, in the long-term, as reported in the revised article, differences that were not statistically significant were found, as the fixed time-point estimates of mortality at 1 year corrected age (imputing loss to follow-up as a death) were 19.4% for lucinactant and 24.2% for poractant (odds ratio: 0.64 [95% confidence interval: 0.32–1.27]). Of course, these confidence intervals are wide and do not exclude a benefit for lucinactant, but they also do not exclude a possible benefit for poractant. In other words, the results in terms of death at 1 year were inconclusive.

Furthermore, the choice of the Wilcoxon test is unusual, because the default is the log-rank test. This choice of test is particularly questionable if there is neither evidence that the use of the Wilcoxon test was specified beforehand (eg, at the time of the study plan) nor scientific justification for that choice.

As the authors themselves claimed in their discussion, the Wilcoxon test emphasizes earlier treatment differences, and it is more powerful when early differences are larger than late differences. As highlighted above, this approach does not address the main clinical message and, in any case, such gain in power is illegitimate if the choice of test is based on an inspection of the data.

In short, we believe that the conclusions drawn by the authors exceed what are justified by the STAR study2 results and seem to be in contrast with the considerations that the same authors raised in their original article, in which they stated that the 2 surfactants were equally effective in preventing respiratory distress syndrome and had comparable incidences of common complications, which suggests similar efficacy and safety.

Stephen Senn, PhD
Department of Statistics
University of Glasgow
Glasgow G12 8QQ, United Kingdom

Marco Costantini, MSc
Giada Rizzi, BSc
Biometrics Unit
Chiesi Farmaceutici SpA
43100 Parma, Italy

Financial Disclosure: Professor Senn acts as a consultant for Chiesi Farmaceutici SpA, and Mr Constantini and Ms Rizzi are employees of Chiesi Farmaceutici SpA.
REFERENCES

In Reply.—

In our article,1 we reported the 1-year follow-up of preterm infants enrolled in 2 contemporaneous clinical trials that compared lcinucant, a peptide-containing synthetic surfactant, with colfosceril and animal-derived surfactants. The primary efficacy end points for these trials were described in their original publications.2,3 The primary conclusion from this 1-year follow-up was that lcinucant is at least as good as, if not superior to, animal-derived surfactants for prevention of respiratory distress syndrome. In response to this report, Lavin and Senn et al express concerns primarily related to our analytical approach to the data.

Any retrospective analysis, be it a pooled patient-level analysis or a meta-analysis, can be biased, and we addressed the limitations of previously published meta-analyses4,5 in our article. Although variation in birth weight, gestational age, prenatal steroid use, and Apgar score in previously published meta-analyses is both substantial and clinically relevant, in our report these differences were slight and clinically unimportant. These trivial differences should reduce concerns about the validity of our report.

We clearly described that the 1-year follow-up in both trials was planned before initiation and that all outcomes were prespecified. The study investigators remained blinded to 1 year corrected age. The aim of our article was to report those findings, the first such report among surfactant-comparison trials.

The analytical approach in our 1-year follow-up was also prespecified, and additional analyses performed beyond the prespecified plan were clearly delineated in the article to avoid misleading the reader. Indeed, our approach would satisfy the rigorous regulatory standard quoted by Senn et al. However, publication in top-quality peer-review journals is the clinician’s gold standard for guiding treatment decisions. Given the relevance and impact factor of Pediatrics, which also published the lcinucant trials, our reporting of these 1-year outcomes has met all standards.

The time points chosen for the analyses were clinically relevant. The first 28 days after birth defines the newborn period, the saccular phase of normal lung development concludes at 36 weeks’ gestational age, and infancy concludes at 1 year of age. Clinically, it is most helpful to present data to neonatologists in a way that is consistent with their considerations of the patient population for which they care. The 1-year corrected age time point provided a concrete and universally understood clinical milestone while also providing a reasonable period of follow-up when relative differences in the degree of lung health and overall well-being may affect survival. The fact that there were minimal (<2%) losses to follow-up at 1 year corrected age in both trials gives us substantial strength to affirm our conclusions.

Interestingly, after criticizing our analyses, both Lavin and Senn et al provide their own retrospective interpretation, contradicting their early admonitions of such analyses to draw conclusions. In addition, Lavin suggests a pooled analysis, which requires more assumptions of the data than a meta-analysis. As suggested by Lavin and Senn et al, a logistic regression may be suitable when clinical milestones, such as 28-day survival, create landmark analyses. With a longer-term observation period such as a 1-year follow-up, a more sensitive and powerful time-to-event analysis such as survival analysis is the analysis of choice, even by regulatory agencies. Given patient-population characteristics in our studies, we used analytical methods that avoid bias in either direction and are most sensitive for identifying significant effects. The majority of deaths occur within 28 days after birth. Thereafter, the number of deaths out to 1 year corrected age is modest. Therefore, the Wilcoxon test provided, as Senn et al assert, an analysis that is sensitive to short-term differences and avoids bias on the basis of the small changes after the first 28 days after birth.

Regarding the figures, we regret that there was an error in the original publication; however, a correction was subsequently published to correct this misprint.

Statistical analyses of data from clinical studies are a means to understanding such data. Clinicians need to understand the degree to which study results can be applied to their scope of practice, and statistics allow them to judge whether these results are beyond mere random observations. In our 1-year follow-up, a numerically superior survival rate was observed for infants who were treated with lcinucant compared with those treated with beractant, poractant alfa, or colfosceril palmitate, with no difference in morbidity despite the proportionally higher number of survivors in the groups.
treated with lucinactant. Discussion of alternative statistical methods will not change this observation. Clinicians at the bedside are now left to decide how the results will affect their practice.

Fernando Moya, MD
Department of Neonatology
Coastal Area Health Education Center
Wilmington, NC 28402

Sunil Sinha, MD, PhD, FRCP, FRCPC
Department of Pediatrics and Neonatal Medicine
University of Durham and James Cook University Hospital
Middlesbrough TS4 3BN, United Kingdom

Ralph B. D’Agostino, Sr
Framingham Study: Statistics/Data Management
Department of Mathematics/Statistics and Public Health
Boston University
Boston, MA 02215

REFERENCES
doi:10.1542/peds.2007-2592

Relevance of Motor Skill Problems in Victims of Bullying

To the Editor.—

We were pleased to see that bullying again was discussed in the July issue of Pediatrics.1 The authors showed that bully victimization is associated with the prevalence of headache, stomachache, sleeping problems, nervousness, and more frequent use of medication. Previous longitudinal studies have confirmed that bullying leads to increased risk of physical symptoms, whereas the direction of causality for psychological factors remains uncertain. Unfortunately, the results from this study shed no additional light on this complex issue. Although the directionality of bullying and emotional problems is complicated, it is, according to our view, essential to prevent bullying. As adult psychiatrists we know that exposure to severe bullying in childhood will continue to influence the person’s self-image and trust in others in adulthood. Also, those who have only been afflicted by milder forms of bullying will probably remember this humiliation all through their lives.

A relationship between bully victimization and having poor social skills has been shown,2,3 but to determine if poor social skills are the result of being bullied or are the cause requires not only large prospective longitudinal studies but also time-consuming assessments. Other ways to disentangle this complex matter could be to study conditions related to the target issue, but where the direction of causality is less disputed. Children with autism spectrum disorder (ASD) run a high risk for becoming severely bullied, as testified to in numerous autobiographies. Indisputably, it is more likely that the subjects’ poor social skills, defined in the diagnostic criteria for ASD, give rise to the bullying than that bullying causes their ASD.

According to Rutter,4 autism might be based on a continuously distributed dimension rather than a distinct category. A broader phenotype of milder varieties of ASD is probably common in the general population. If poor social skills frequently result in being bullied, subjects with the ASD phenotype would be overrepresented in the bully-victim population. One way of identifying this phenotype could be to study motor skills. Motor impairment is strikingly common in ASD.5,6 We assume that children with motor impairment would perform poorly in physical education. On the basis of these assumptions, we hypothesized that bullying victimization would be more common for children with poor performance in physical education compared with those with superior performance.

To challenge our hypothesis we conducted a somewhat unorthodox study on the audiences, all professionals, attending a series of day conferences on psychiatric disorders in adults during 2006–2007 (N = 1043; 85% female; mean ± SD age: 44 ± 11 years; range: 19–71 years). The conferences were held in 7 different Swedish cities on 9 occasions. During the course of the day the participants were asked to respond anonymously to a few questions, presented 1 at a time, in a slide show. They were asked if they had been bullied in school (not at all, a little, or severely) and to estimate their own performance in physical education. We supplemented our investigation on the 3 last conferences by adding an additional question on their academic performance in school (above average, average, or below average). Bullying was not otherwise mentioned in the conference lectures until after this assignment was accomplished.

Exposure to bullying was equally common at all locations and in both genders. Thirty percent reported that they were bullied in school at least to some degree. More men than women reported superior performance in
physical education ($\chi^2 = 10.1$; degrees of freedom [$df = 2$; $P = .006$). Academic performance, however, did not differ between genders. Those participants who acknowledged being poor performers in physical education (27% [$n = 263$]) were exposed more often to bullying than those who reported superior performance ($\chi^2 = 28$; $df = 2$; $P = .000001$), whereas the relationship between reported academic achievement and exposure to bullying was nonsignificant ($n = 316$; $\chi^2 = 5.2$; $df = 2$; $P = .08$).

In our view, being a target of bullying per se does not determine a poor psychosocial outcome. We propose that a poor psychosocial outcome in bullied subjects mostly reflects a previous low social capacity that may generate aversive feelings in peers who instinctively respond by bullying. This is likely to result in a vicious cycle with increased anxiety and maladjustment, which possibly explains more frequent use of medication. However, one should keep in mind that having an ASD phenotype, per se, may increase the risk for several physical and mental symptoms.

To prevent bullying it is crucial to understand why certain children are at risk for becoming victims and to enhance other children’s understanding and acceptance for those with poor social skills. To accomplish this is a true challenge, but without adequate guidance it is difficult for any child to suppress instinctive behaviors, and there is a high cost for those who become their targets. Political correctness says that being a victim of severe bullying is likely to be a random event. Everyone who has gone through school knows that this is not true, which should be kept in mind when planning intervention and prevention programs against bullying.

Susanne Bejerot, MD
Department of Clinical Neuroscience
Section Psychiatry St Göran
Karolinska Institute
SE-112 81 Stockholm, Sweden

Mats Humble, MD
Department of Clinical Sciences
Lund University
SE-205 02 Malmö, Sweden

Biological Psychiatry Research Group
Univestitetssjukhuset MAS
SE-205 02 Malmö, Sweden

REFERENCES

doi:10.1542/peds.2007-2187

In Reply.—

By definition, bullying takes place among children with different status levels, from children with higher status toward children with lower status. We agree that social skills are crucially important for adolescents to achieve status in the group and that, consequently, lack of social skills, also when deriving from autism spectrum disorder (ASD), will lead to lower status in the group and increase the child’s risk of bullying victimization. Although we agree that ASD may increase risk of bullying victimization for the individual, it is unlikely that ASD explains most of the prevalence of bullying victimization at the population level in countries with a high prevalence of bullying (eg, Denmark).

The prevalence of bullying among Danish children was 25% in 1994 and 25% in 1998. Four years later, after instituting intensive national-level policies directed at bullying in Danish schools, the prevalence was 11%. The Danish interventions were directed at the school and school-class environments and were highly effective in lowering the prevalence of bullying. Also, we found large variations in the prevalence of bullying victimization between schools and between countries: in Danish schools in 1998 ranging between 10% and 46% and in our international study ranging between 6.3% (Swedish girls) and 41.4% (Lithuanian boys). It is unlikely that these prevalence differences over time and between schools and countries are explained by ASD prevalence differences.

Politically correct or not, bullying is not a random event. It occurs more often toward children who are fragile for one reason or another. This means that strengthening any child’s resources, social as well as mental and physical, is an important step to reduce each individual’s risk of being bullied. However, in populations with a high prevalence of bullying, the overall risk of bullying at the population level is not likely to be attributable to the prevalence of children with ASD or other “weaknesses” but, rather, likely to be largely attributable to the acceptance of bullying behavior in the environment. Only this explains the large prevalence differences between countries, schools, and classes, and
only this explains that the prevalence of bullying victimization in Denmark can be reduced by 56% over a period of 4 years.

The above-mentioned results, as well as international literature, state that intervention against bullying is effective when it addresses the school and school-class environment.5–7 In countries such as Sweden with very low bullying prevalence,4 environmental preventive strategies may have fully served their purpose, and individual-level strategies may be more important in reaching an absolute level of 0% prevalence of bullying.

REFERENCES

1. Olweus D. Bullying at School: What We Know and What We Can Do. Oxford, United Kingdom: Blackwell; 1993
doi:10.1542/peds.2007-2569

Keith K. Lau, MDa, Jayson M. Stoffman, MDb, Suzane Williams, MDc, Patricia McCusker, MDd, Leonardo Brandao, MDc, Sanjay Patel, MScd, Anthony K. C. Chan, MDae, for the Canadian Pediatric Thrombosis and Hemostasis Network

Department of Pediatrics, University of California, Davis, California; bDepartment of Pediatrics, University of Manitoba, Winnipeg, Manitoba, Canada; cDepartment of Pediatrics, University of Toronto, Toronto, Ontario, Canada; dDivision of Hematology/Oncology, Department of Pediatrics, McMaster University, Hamilton, Ontario, Canada

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

Renal vein thrombosis is a complication that occurs in neonates with various underlying risk factors. It carries a grave prognosis for affected kidneys. Anticoagulant and fibrinolytic therapies have been promoted in the past with anecdotal success in some circumstances. However, prospective controlled trials are still lacking, and to date there have been no evidence-based guidelines available for the treatment of neonates with renal vein thrombosis. We retrospectively reviewed all the available medical literature pertaining to renal vein thrombosis published in English during the past 15 years. A total of 271 patients from 13 case series were identified by using the terms “renal vein thrombosis” and “neonates” via PubMed and Cochrane Library searches. Data then were extracted from each of the studies for analysis. During the past 15 years, a male predominance (67.2%) in neonatal renal vein thrombosis has been reported. More than 70% of patients had unilateral renal vein thrombosis, which was more prevalent on the left side (63.6%). The thrombus involved the inferior vena cava and was associated with adrenal hemorrhage in 43.7% and 14.8% of neonates, respectively. Forty percent of the patients were treated conservatively with supportive care alone. Among those patients who received anticoagulation therapy, unfractionated heparin and low molecular weight heparin were used alone in 21.6% and 20.7% of the patients, respectively. Fibrinolytic treatment alone was used in 11.2% of the patients. Only a minority of patients were treated with antithrombin (1.7%), warfarin alone, (0.9%) or underwent surgical intervention (0.3%). The majority (70.6%) of the involved kidneys became atrophic. A total of 9 neonates died with non–renal vein thrombosis–related conditions during the study period. Evidence-based recommendations on treatment cannot be made at the present time. Cooperative prospective studies that involve multiple centers are needed to elucidate the optimal treatment for neonatal renal vein thrombosis.
The exact incidence of renal vein thrombosis (RVT) in neonates is difficult to determine, because large-scale population-based epidemiologic studies are lacking. It has been reported to occur in 0.5 per 1000 admissions to NICUs according to a large international registry. On the other hand, a recent study from Germany suggested that the incidence of symptomatic RVT in neonates was at least 2.2 per 100 000 live births. Since the systematic and detailed review on the hemostatic complications of pediatric renal disease by Andrew and Brooker more than 10 years ago, multiple case series and reports on neonates with RVT have been published. We reviewed the available English-language medical literature pertaining to neonatal RVT over the past 15 years to analyze and examine the data that are currently available on this neonatal disease. We intended to evaluate the efficacy of our current management strategies and identify potential improvement in the care of neonates with RVT.

METHODS
We searched the PubMed database from the National Library of Medicine using the keywords “renal vein thrombosis” and “neonates” with the limits set to only English-language articles and those that involved human subjects. An additional search was performed via the Cochrane Central Register of Controlled Trials and the Cochrane database of systemic reviews (Issue 4, 2006). We included all case reports and case series that were published in English from January 1992 to December 2006 that reported neonates with RVT. We excluded studies that contained discussion related to neonatal RVT but did not report on an actual case and those reports with 2 or fewer patients. Because the diagnostic criteria of RVT were not clearly stated in every study, we could not scrutinize the accuracy of the diagnosis of each reported case.

Data were extracted from individual studies onto a spreadsheet (Excel; Microsoft, Redmond, WA) for additional analysis. Because the design and methodologies used varied among the studies, some did not have all the information that we intended to analyze. Data are included in the analysis if they were actually described in the study; if a test was not mentioned in the study, the result was entered as unknown.

RESULTS
The search strategies identified 77 publications; 64 were excluded from analysis. Thirty-seven of those reports were excluded because the number of cases included were 2 or fewer. One case series, reported by the 1-800-NO-CLOTS registry, was also excluded from the study because of the potential overlap of patients from other reported cases. Another case series of 23 children with RVT was also excluded because 4 of the patients presented after the neonatal period but could not be removed from the aggregate analysis. As a result, a total of 13 case series were included for analysis. All of the reported cases had objective documentation of RVT from ultrasound.

Patient Characteristics
Table 1 describes the studies that were included in our review. A total of 271 neonates were included, with a male predominance (67.2% [127 of 189]). Figure 1 depicts the time of onset of RVT. Although all the patients were diagnosed to have RVT within the first month of life, detailed information on the exact time of disease onset was not available from all studies. However, data compiled from studies with that information show that 7.3% (6 of 82), 67.1% (55 of 82), and 25.6% (21 of 82) of the neonates presented in utero, within 3 days, and more than 3 days after birth, respectively. Most of the patients were born at term (71.3% [97 of 136]).

Presentation
Of the neonates, 70.3% (173 of 246) had unilateral RVT; 63.6% (63 of 99) of these cases involved the left kidney.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>n</th>
<th>F</th>
<th>M</th>
<th>Unilateral</th>
<th>Bilateral</th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messinger et al. (2006)</td>
<td>28</td>
<td>8</td>
<td>20</td>
<td>25</td>
<td>3</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Winyard et al. (2006)</td>
<td>23</td>
<td>6</td>
<td>17</td>
<td>10</td>
<td>13</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Marks et al. (2005)</td>
<td>43</td>
<td>15</td>
<td>28</td>
<td>24</td>
<td>19</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Proesmans et al. (2005)</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Kosch et al. (2004)</td>
<td>59</td>
<td>24</td>
<td>35</td>
<td>43</td>
<td>16</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td>Heller et al. (2000)</td>
<td>25</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Bökenkamp et al. (2000)</td>
<td>35</td>
<td>NA</td>
<td>NA</td>
<td>28</td>
<td>7</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Schmidt and Andrew (1995)</td>
<td>21</td>
<td>5</td>
<td>16</td>
<td>18</td>
<td>3</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Keiden et al. (1994)</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Huss et al. (1994)</td>
<td>6</td>
<td>NA</td>
<td>NA</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Laplante et al. (1993)</td>
<td>8</td>
<td>NA</td>
<td>NA</td>
<td>7</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Orazi et al. (1993)</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Nowak-Gött et al. (1992)</td>
<td>8</td>
<td>NA</td>
<td>NA</td>
<td>4</td>
<td>4</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Total</td>
<td>271</td>
<td>62</td>
<td>127</td>
<td>173</td>
<td>73</td>
<td>63</td>
<td>36</td>
</tr>
</tbody>
</table>

F indicates female; M, male; L, left kidney; R, right kidney; NA, not applicable.
Figure 2 depicts the features of the patients at presentation. The thrombus extended into the inferior vena cava in 43.7% (90 of 206) of the patients. Adrenal hemorrhage was present in 14.8% (31 of 210) of the patients. Because most of the patients were reported as case series rather than as individual cases, individual clinical features at presentation could not be analyzed and correlated with outcomes. Perinatal risk factors including asphyxia were identified in 31.9% (69 of 216) of the reported cases. Other known risk factors such as maternal diabetes mellitus and dehydration were reported in 8.1% (17 of 210) and 1.5% (3 of 206) of the patients, respectively. Although all the patients had ultrasound-documented RVT, most of them also had at least 1 of the 3 cardinal signs of RVT at presentation: 56.2% (73 of 130) had macroscopic hematuria, 45.4% (83 of 183) had a palpable abdominal mass, and 47.5% (85 of 179) had thrombocytopenia.

Prothrombotic Risk Factors
Not all of the studies collected data on their patients’ prothrombotic risks. Before 1996, the only prothrombotic abnormality reported was the presence of the lupus anticoagulant. Since that time, other prothrombotic factors, including protein C, protein S, plasma antithrombin III activity, lipoprotein(a), factor V Leiden mutation, prothrombin gene mutation, and methylenetetrahydrofolate (MTHFR) thermolabile mutation have also been screened in some studies. Among patients with RVT in whom prothrombotic factors were investigated, 53% (79 of 149) had at least 1 risk factor identified. The recurrence rate of thromboembolic events was reported in 1 study. \^ With a median follow-up time of 4 years (range: 0.6–15 years), 6.8% (4 of 59) of the patients with RVT had a second thrombotic event, occurring during puberty in 3 of the 4 patients. All 4 patients had at least 1 prothrombotic risk factor.

Treatment Modalities
Ten of the 13 studies reviewed contained data on treatment of RVT. Figure 3 illustrates the different treatment modalities that the patients received: 39.7% (92 of 232) received supportive therapy alone, and 21.6% (50 of 232) and 20.7% (48 of 232) of the patients received unfractionated heparin or low molecular weight heparin (LMWH) alone, respectively. Slightly more than 11% (11.2% [26 of 232]) were given fibrinolytic treatment. A small minority of patients were managed with antithrombin (1.7% [4 of 232]) or warfarin (0.9% [2 of 232]) alone or surgical interventions (0.3% [9 of 232]). Approximately 4% (3.9% [9 of 232]) of the patients received combinations of treatment that included heparin, fibrinolytic agents, warfarin, and protein C concentrate.

Outcomes
Renal outcomes were not reported consistently in the studies. Regardless of the treatment received, irreversible damage was found in 70.6% (156 of 221) of the affected kidneys at last follow-up. In patients treated with unfractionated heparin/LMWH or supportive care, 75.3% (61 of 81) and 72.5% (45 of 62), respectively, of the affected kidneys were found to be atrophic at last follow-up (Table 2). Although the definitions used for
chronic renal insufficiency were heterogeneous and often unspecified, 8 neonates (3.0% [8 of 271]) required chronic renal replacement therapies and renal transplantations. Four of these 8 neonates had bilateral RVT; the extent was not documented for the other 4 patients.

Seven of the 13 studies provided follow-up information for hypertension: 19.3% (27 of 140) of the patients had persistent elevation of blood pressure. Laterality of the involved kidneys was provided for 120 patients: 18.9% (14 of 74) of the neonates with unilateral RVT and 21.7% (10 of 46) of the neonates with bilateral RVT had persistent elevated blood pressure (Table 3).

A total of 9 deaths were reported, all of which were related to coexisting medical conditions such as respiratory failure in 3 neonates, multiorgan failure in 1 neonate, and sepsis/meningitis in another neonate. The cause of death of the other neonates was not provided by the authors of the report.

D I S C U S S I O N

RVT is the most common non–catheter-related thrombosis in infancy and occurs primarily in the newborn period.1 Although many reviews of RVT can be identified in the literature, evidence-based management guidelines are still lacking. In this study, we retrospectively analyzed all the available case series with more than 2 patients on neonatal RVT published in English over the last 15 years. Our intent was to gather the accumulated experience from different sources, evaluate the current management strategies for neonatal RVT, and provide direction for future research in this important hemostatic disorder in neonates.

In an earlier review that included 268 children with RVT, of which 212 were neonates, the incidence in boys and girls was similar and the left and right sides were affected equally.3 Other studies have noted a male predominance, as we have observed in this review. It is uncertain whether this represents a changing demographic of RVT. In our study, the incidence of bilateral RVT was 29.7%, compared with 24% in the earlier review. Preterm infants have been considered to be at risk of developing RVT, and we found that 28.7% of the affected neonates were born before 36 weeks of gestation. Although the classical “triad” of RVT (macroscopic hematuria, palpable abdominal mass, and thrombocytopenia) has been well described, these 3 elements are not always found at presentation. Winyard et al44 recently showed that only 23.5% (5 of 21) of their neonates had the triad at presentation. Perinatal asphyxia, dehydration, and maternal diabetes mellitus are established risk factors for RVT. However, we found that less than one third of the affected neonates had a history of perinatal asphyxia, and dehydration and maternal diabetes mellitus were even less common.

Hereditary prothrombotic risk factors may also play a role in the pathogenesis of neonatal RVT, and routine screening for known procoagulant abnormalities has been suggested.24,45 We found that 53% (79 of 149) of the patients who had been investigated were found to have 1 or more prothrombotic risk factors. In a recent study of 301 children with various types of spontaneous thromboembolism, 21.3% of the patients experienced a recurrence at a median of 3.5 years after cessation of prophylaxis.54 Among those children whose symptoms recurred, 48.4% had a single prothrombotic risk factor and 46.9% had 2 or more risk factors. The authors concluded, therefore, that it is reasonable to screen children with symptomatic thromboembolic diseases for prothrombotic risk factors. However, 1 study in our review47 found that only 6.8% (4 of 59) of the affected neonates had a recurrence of thrombosis, and 3 of the 4 patients had the recurrence during puberty. Another study did not find any recurrence of the thrombosis beyond the neonatal period in neonates with prothrombotic risk factor at a median follow-up of 3.7 years (range: 0.5–20.2 years).45 Because there are no data to indicate that routine screening in neonates with RVT reduces the recurrence risk, there is no evidence to support routine screening for prothrombotic risk factors in neonates with RVT.

Ultrasound is a useful and convenient clinical tool for

---

**TABLE 2**

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Heparin Treated</th>
<th>Supportive Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keiden et al (1994)</td>
<td>2 (2)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>Nuss et al (1994)</td>
<td>3 (6)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Orazi et al (1993)</td>
<td>0 (0)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>Total</td>
<td>61 (81)</td>
<td>45 (62)</td>
</tr>
</tbody>
</table>

The incidence is shown as number of neonates with atrophic kidneys and the total number of patients treated in parentheses.

---

**TABLE 3**

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Unilateral</th>
<th>Bilateral</th>
<th>Nonspecified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winyard et al (2006)</td>
<td>1 (10)</td>
<td>3 (13)</td>
<td>—</td>
</tr>
<tr>
<td>Marks et al (2005)</td>
<td>1 (10)</td>
<td>2 (12)</td>
<td>—</td>
</tr>
<tr>
<td>Kosch et al (2004)</td>
<td>10 (43)</td>
<td>3 (16)</td>
<td>—</td>
</tr>
<tr>
<td>Bökenkamp et al (2000)</td>
<td>—</td>
<td>—</td>
<td>3 (20)</td>
</tr>
<tr>
<td>Keiden et al (1994)</td>
<td>0 (5)</td>
<td>0 (1)</td>
<td>—</td>
</tr>
<tr>
<td>Nuss et al (1994)</td>
<td>0 (3)</td>
<td>0 (1)</td>
<td>—</td>
</tr>
<tr>
<td>Orazi et al (1993)</td>
<td>2 (3)</td>
<td>0 (1)</td>
<td>—</td>
</tr>
<tr>
<td>Total</td>
<td>14 (74)</td>
<td>10 (46)</td>
<td>—</td>
</tr>
</tbody>
</table>

The incidence is shown as number of neonates with elevated blood pressure and the total number of patients with unilateral or bilateral RVT in parentheses. — indicates that data were not available.
the diagnosis of RVT. Sonographic features of RVT include enlarged and echogenic kidneys with attenuation or loss of corticomedullary differentiation. Calcification and thrombus may be seen extending outside the kidneys to the inferior vena cava. Ultrasound may also be of prognostic significance for RVT. Winyard et al44 reported a negative correlation between renal length and renal outcomes. Doppler studies are particularly useful for detecting the resistance or absence of flow in renal venous branches and collateral vessels. Although the blood flow in the main renal vein and its branches may be normal, there may be an increase in resistance in the renal arteries caused by thrombosis in the small intrarenal veins. Renal scarring and atrophy are well-recognized features after RVT in affected kidneys, which can also be assessed by a radionuclide scan.

Heparin therapy has been proposed for neonates with RVT, especially in those with bilateral RVT and with inferior vena cava involvement. Recently, the trend has been toward the use of LMWH rather than unfractionated heparin. However, almost 40% of the affected neonates were managed by supportive care alone. Our data show that the treatment of neonates with RVT varies greatly among centers, and evidence-based recommendations on this practice are not available. The data suggest that the renal outcomes were similar between supportive treatment and heparin therapies, and a similar proportion of affected kidneys became atrophic in neonates who were managed supportively or with heparin. RVT in neonates often leads to irreversible damage, and anticoagulant therapies may not have an impact on the long-term outcomes.

Although there is no consensus on the management of neonatal RVT, the management should involve a multidisciplinary team that includes neonatologists, radiologists, hematologists, and nephrologists. During the acute phase, supportive management of electrolytes and fluid balance with input from the nephrologists is essential, particularly if the neonates are in acute renal failure and require renal replacement therapies. The hematologists should be consulted on whether and when the neonates require anticoagulation or fibrinolytic treatments. The response to the treatment should be monitored by the clinical team in conjunction with the radiologists and ultrasonographers.

Affected neonates must be followed closely for renal complications such as hypertension, atrophy, functional loss, and chronic renal insufficiency. In our review, reporting on outcomes of the affected neonates was incomplete in many of the publications; the length of follow-up and definitions for chronic kidney diseases were also variable. Persistent hypertension was reported in one fifth of the neonates, and 8 neonates (3.0%) were reported to need chronic dialysis or kidney transplantation. Because of the heterogeneity of treatment approaches and outcome measures, we are not able to draw any conclusion on the impact of different treatment strategies on long-term outcomes.

Nine deaths were reported, and all were non–RVT related. The mortality rate in children with thromboembolic disease from all causes was reported as 5% in a large international registry1 that was published more than 10 years ago. The mortality rate of neonates with RVT that we observed in our study was 3.3% (9 of 271).

Although we retrospectively pooled the available data on neonates with RVT over the last 15 years from multiple centers, we still cannot make any evidence-based recommendations for current treatment strategies. Randomized, controlled trials may not be feasible because of the low incidence of RVT and limited patient numbers in individual centers. However, a prospective multicenter trial with standardized protocols for treatment and long-term follow-up of neonatal RVT should be possible and practical.

CONCLUSIONS

Our study illustrates the variability of treatment strategies for neonates with RVT and, potentially, a change in demographics of this condition. Renal outcomes of the affected kidneys are still unsatisfactory for the majority of neonates with RVT. Evidence-based recommendations for the optimal treatment of neonatal RVT are not possible, but prospective multicenter trials with standardized protocols are feasible and should be undertaken as soon as possible to identify the most appropriate treatment strategies for neonates with RVT.

REFERENCES

10. Weinschenk N, Pediris M, Fiascone J. Combination thrombo-


Interpreting and Managing Blood Lead Levels of Less Than 10 μg/dL in Children and Reducing Childhood Exposure to Lead: Recommendations of the Centers for Disease Control and Prevention Advisory Committee on Childhood Lead Poisoning Prevention

Helen J. Binns, MD, Carla Campbell, MD, Mary Jean Brown, ScD, RN, for the Advisory Committee on Childhood Lead Poisoning Prevention

a Feinberg School of Medicine, Northwestern University, Chicago, Illinois; bChildren’s Hospital of Philadelphia, Philadelphia, Pennsylvania; c Division of Environmental and Emergency Health Services, National Center for Environmental Health/Agency for Toxic Substances and Disease Registry, Center for Disease Control and Prevention, Atlanta, Georgia

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

Lead is a common environmental contaminant. Lead exposure is a preventable risk that exists in all areas of the United States. In children, lead is associated with impaired cognitive, motor, behavioral, and physical abilities. In 1991, the Centers for Disease Control and Prevention defined the blood lead level that should prompt public health actions as 10 μg/dL. Concurrently, the Centers for Disease Control and Prevention also recognized that a blood lead level of 10 μg/dL did not define a threshold for the harmful effects of lead. Research conducted since 1991 has strengthened the evidence that children’s physical and mental development can be affected at blood lead levels of <10 μg/dL. In this report we provide information to help clinicians understand blood lead levels < 10 μg/dL, identify gaps in knowledge concerning lead levels in this range, and outline strategies to reduce childhood exposures to lead. We also summarize scientific data relevant to counseling, blood lead screening, and lead-exposure risk assessment. To aid in the interpretation of blood lead levels, clinicians should understand the laboratory error range for blood lead values and, if possible, select a laboratory that achieves routine performance within ±2 μg/dL. Clinicians should obtain an environmental history on all children they examine, provide families with lead-prevention counseling, and follow blood lead screening recommendations established for their areas. As circumstances permit, clinicians should consider referral to developmental programs for children at high risk for exposure to lead and more frequent rescreening of children with blood lead levels approaching 10 μg/dL. In addition, clinicians should direct parents to agencies and sources of information that will help them establish a lead-safe environment for their children. For these preventive strategies to succeed, partnerships between health care providers, families, and local public health and housing programs should be strengthened.
LEAD IS A common environmental contaminant, and exposure to lead is a preventable risk in all areas of the United States. Lead is associated with negative outcomes for children, including impaired cognitive, motor, behavioral, and physical abilities. In 1991, the Centers for Disease Control and Prevention (CDC) defined the blood lead level (BLL) that should prompt public health actions as 10 μg/dL. Concurrently, the CDC recognized that a BLL of 10 μg/dL did not define a threshold for the harmful effects of lead. Research conducted since 1991 has strengthened the evidence that the physical and mental development of children can be affected at BLLs of <10 μg/dL.

In 2002 to 2004, a workgroup of the CDC Advisory Committee on Childhood Lead Poisoning Prevention (ACCLPP) reviewed the scientific literature regarding adverse health effects associated with BLLs of <10 μg/dL, including 23 published reports that analyzed 16 separate populations with IQ or general cognitive index outcomes and 12 publications related to other health outcomes. In its 2005 report, the workgroup concluded that an inverse association exists between BLLs and cognitive function, with no evidence of a weaker association in populations with lower BLLs. The direct evidence for this inverse association was strongest in a study conducted in Rochester, New York, which included children who were born in 1994 or 1995, enrolled at 6 months of age, and monitored for 5 years. The majority of children studied had BLLs of <10 μg/dL throughout the study period. The IQ/BLL relationship was described most accurately by a nonlinear negative association, with a decrease in IQ of >7 points over the first 10 μg/dL increase in lifetime average BLL. On the basis of the evidence, the workgroup concluded that a causal association between lead exposure and impaired cognitive functioning was most likely. However, the potential for residual confounding, particularly by social factors, made the strength and shape (ie, linear or nonlinear) of this association across BLLs uncertain. In addition, the workgroup concluded that children with BLLs of <10 μg/dL should not be classified as “lead poisoned.” The report noted that no safe BLL in children has been identified.

Two studies published subsequently reported negative effects of BLLs of <10 μg/dL on developmental outcomes. One study, which included participants from the Rochester cohort and from 6 other prospective studies of children with peak BLLs across a range of values, reaffirmed an inverse association between low BLLs and IQ. Those studies accounted for key potential confounders, including maternal IQ. Home Observation for Measurement of the Environment Inventory score (which is a measure of the quality and quantity of stimulation and support available to a child in the home environment), maternal education, and birth weight.

Although the ACCLPP previously reviewed case management for children with BLLs of ≥10 μg/dL, this is the first ACCLPP report to summarize scientific information relevant to clinical management for children with BLLs of <10 μg/dL. This report also outlines recommendations from the ACCLPP to reduce childhood exposure to lead. Information on assessments of environmental history and prevention strategies to decrease exposure to lead was published previously and is not included in this report.

METHODS
The ACCLPP provides advice and guidance to the US Department of Health and Human Services and the CDC regarding new scientific knowledge and technological developments and their practical implications for preventing childhood lead poisoning and recommends improvements as needed. ACCLPP members are selected on the basis of their expertise in childhood lead poisoning prevention, screening, diagnosis, and medical management. ACCLPP liaisons represent federal agencies and organizations with particular interest and expertise in childhood lead poisoning prevention.

In October 2003, the ACCLPP formed another workgroup, consisting of 3 pediatricians and a CDC health scientist, to review the scientific literature regarding clinical management options for BLLs of <10 μg/dL and to outline recommendations for clinical care providers. On the basis of its analysis, the workgroup developed draft recommendations that were reviewed and then adopted by the ACCLPP in February 2006.

RESULTS

Historic Trends in Children’s BLLs in the United States
Since 1976, BLLs in US children 1 to 5 years of age have decreased substantially (Table 1), primarily as a result of policies that have reduced the dispersal of lead into the environment. However, many US children continue to be exposed to lead, primarily in their homes. Overt clinical symptoms of lead intoxication are uncommon in the United States, and lead evaluation and management strategies typically are intended to reduce the negative effects of lead on central nervous system development in children who are clinically asymptomatic. Because no safe BLL has been defined, small reductions in population-level exposures to lead would likely affect substantial numbers of children and could be expected to reduce

<table>
<thead>
<tr>
<th>Year</th>
<th>Proportion With BLLs of ≥10 μg/dL, %</th>
<th>Geometric Mean BLL, μg/dL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976–1980</td>
<td>88.2</td>
<td>15.0</td>
</tr>
<tr>
<td>1999–2002</td>
<td>1.6</td>
<td>1.9</td>
</tr>
</tbody>
</table>
the number of children with adverse health outcomes associated with lead exposure.14

**BLL Measurements**

As with any biological test, BLL measurements have inherent uncertainties resulting from imprecise analytic techniques and preanalytic variables (eg, the specimen collection process). However, the imprecision/measurement value ratio, particularly at BLLs of <10 μg/dL, is relatively high. The degree of inherent error in BLL analyses varies according to the analytic method used, whichever method is used, laboratory performance depends on the procedures and skills of the laboratory team.15,16 Federal regulations allow laboratories that perform BLL testing to operate with a total allowable error of ±4 μg/dL or ±10%, whichever is greater. Consequently, at BLLs of ≤10 μg/dL, a laboratory might operate within an error range of 8 μg/dL and still meet federal proficiency standards. For example, an actual BLL of 7 μg/dL could be reported as any value ranging from 3 to 11 μg/dL and still remain within the allowable error limit. A study of duplicate testing of identical blood samples (all with mean BLLs of <10 μg/dL) at 8 laboratories reported all results as <10 μg/dL and within 3 μg/dL of the overall mean for that specimen value.17 A study conducted in 2001 indicated that the majority of BLL laboratories can achieve routine performance of ±2 μg/dL at concentrations of ≤10 μg/dL without difficulty.18

BLL test reliability also depends on adherence to blood collection techniques that reduce sample contamination. Collection of capillary blood from a finger-stick into a lead-free collection device is an accepted method for obtaining a screening test sample.19-23 and contamination by lead from the skin surface can be minimized if a protocol for proper capillary specimen collection is followed. (A complimentary videotape or DVD, CDC Guidelines for Collecting and Handling Blood Lead Samples: 2004, may be obtained from the National Center for Environmental Health, Division of Laboratory Sciences, Lead and Multielement Proficiency Program.) However, because BLLs determined from capillary blood samples vary from those determined from simultaneously drawn venous samples, elevated capillary BLL results should be confirmed with blood samples drawn through venipuncture. Multiple studies have reported on the uncertainty introduced through collection of capillary blood samples, rather than samples obtained through venipuncture, at thresholds of 10 or 15 μg/dL,19-23 but none has examined the sensitivity or specificity of capillary blood sample collection methods at thresholds of <10 μg/dL.

**Children’s BLL Patterns**

BLLs increase quickly after acute exposure and then gradually (over weeks) reach equilibrium with body stores of lead. Lead is distributed unevenly within the human body; in children, ~70% is stored in the bone compartment.24-26 The residence time of lead in bone can be decades.27 Therefore, elevated BLLs decline within a few weeks to months after acute exposure. However, for children with chronic lead exposure and presumably greater bone lead stores, the decline in BLLs can take much longer.28 Although bone lead levels can provide information regarding past absorption of lead, measurements of lead levels in bone by using x-ray fluorescence instruments are available for research purposes only.

The BLL of a newborn infant reflects closely that of the mother.29 In 1999 to 2002, the geometric mean BLL for US women 20 to 59 years of age was 1.2 μg/dL, with 0.3% having BLLs of ≥10 μg/dL.3 Typically, as infants become more active and increase their environmental exposure, BLLs increase. Longitudinal studies of lead-exposed children have confirmed an increase in BLLs beginning in late infancy, with peak BLLs being reached at 18 to 36 months of age.6,30-32 No studies have examined BLL patterns specifically for children with peak BLLs of <10 μg/dL, although certain studies included children with levels that low. A study of children born in 1994 and 1995, in which >50% of the children had peak BLLs of <10 μg/dL, reported an expected pattern of mean BLLs of 3.4 μg/dL at 6 months of age, 9.7 μg/dL at 24 months of age, and 5.8 μg/dL at 61 months of age.6 A study of children born in Boston, Massachusetts, in 1979 to 1981 identified a mean BLL of 7.2 μg/dL at birth, and subsequent BLLs for those children remained relatively constant (6.2 μg/dL at 6 months of age, 6.8 μg/dL at 24 months of age, and 6.4 μg/dL at 57 months of age).33-35 In both studies, higher levels of lead in home environmental samples were associated directly with higher BLLs in children.34,36 In addition, the Boston study demonstrated an association between the occurrence of home renovation and increased BLLs.34 The BLL patterns for individual children with BLLs of <10 μg/dL vary, depending on environmental exposures.29 More research is needed to understand more thoroughly age-related patterns for BLLs that remain at <10 μg/dL. Even if additional research data become available, however, laboratory uncertainty might interfere with a clinician’s ability to detect patterns for individual children.

Once a high BLL has been established for a child, the time required for the BLL to decline to <10 μg/dL can range from months to years, depending on the duration and dose of exposure. For example, for a group of children who started with BLLs of 10 to 14 μg/dL and received case management services, the mean time required for 50% to achieve BLLs of <10 μg/dL was 9 months.37 The time needed for BLLs of <10 μg/dL to decline in response to interventions is not known.

Multiple studies have confirmed that BLL measurements vary seasonally. For example, a study conducted in Boston reported that BLLs were highest in late June and lowest in March.38 A study performed in Milwaukee,
Wisconsin, indicated that BLLs were higher in the summer than in the winter.39 Some of the variability (higher BLLs in summer) might result from increased exposure to lead in dust and soil in the summer months.40 BLL values for urban children are predicted to be 1 to 2 μg/dL higher in the summer months than in the winter months.41

**Association of BLL Patterns With Developmental Outcomes**

Although BLLs peak in early childhood, when young children are especially vulnerable to lead, negative effects are associated with lead exposure at any age. Multiple studies have examined the effects of lead on children’s developmental outcomes; in those studies, the ages at which BLLs were measured varied, as did the range of ages over which BLLs were averaged.1–4 Statistically significant associations between average BLLs over a specific period (eg, 0–5 years) and various adverse health outcomes have been identified.6,42–44 Other studies reported statistically significant associations with a single lead measurement at a specific age (eg, before birth, at 24 months, or at 6.5 years) or with a peak measurement.6,30,45 Concurrent BLLs (ie, those measured close to the time of neurodevelopmental testing) might demonstrate stronger associations with neurodevelopmental abilities, compared with other BLL measurements.6,8,31,46

Lead has a continuing negative association with IQ as children reach elementary school age. For children who participated in a trial of chelation therapy, a subsequent data analysis indicated that BLLs measured concurrently with developmental testing were associated more closely with children’s cognitive abilities than were peak levels measured at ~2 years of age.47 This association was stronger when children were tested at 7 years of age, compared with 5 years of age, which underscores the continuing need to reduce lead exposure after 5 years of age.

**Strategies to Enhance Children’s Positive Developmental Outcomes**

Although lead is a risk factor for developmental and behavioral problems, its presence does not indicate that these problems will necessarily occur. No characteristic developmental pattern is attributable solely to the effects of lead, and measures of the effects of lead on children are imperfect. For an individual child, neurobehavioral test performance might indicate clinically significant impairments related to lead exposure but might not fully capture the array of negative outcomes caused by lead.14 The effects of lead at levels approaching 10 μg/dL might not be recognizable to the child’s family or clinician and might not be identified through neurobehavioral testing. However, lead exposure might assume greater importance for children with other environmental, genetic, biological, social, or demographic developmental risk factors. The effects of exposure to lead at lower levels might not be evident in testing of individual children and are best evaluated on a community-wide basis.14

Multiple factors influence a child’s development, including how the child is treated by parents and other adult caregivers. The child’s family and personal psychosocial experiences are associated strongly with performance on neurodevelopmental measures and account for a greater proportion of the explained variance in these measures than do BLLs of <10 μg/dL.2,42,44,48 A child’s BLL measurement is estimated to account for 2% to 4% of the variance in neurodevelopmental measures (~4%–8% of the explained variance).2,42,49

All children, regardless of their BLLs, benefit from parental nurturing. For example, a child’s language skills are enhanced by the amount of language addressed to the child (more is better), combined with a predominant pattern of positive feedback.50 This pattern of parenting for children <3 years of age was associated with enhanced language and cognitive skills when children were tested in the third grade.51 Therefore, parents might help counteract the negative effects of lead by providing a nurturing enriched environment during development. Studies examining the effects of lead have attempted to control for this psychosocial factor by including measures such as the Home Observation for Measurement of the Environment Inventory score.7 Although no studies have evaluated specifically the effects of early intervention programs on cognitive or behavioral outcomes in relationship to children’s BLLs, several laboratory studies that applied a nurturing environment to very young animals during lead acquisition demonstrated the beneficial effect of the social environment in ameliorating lead-related negative developmental outcomes.52,53

Early enrichment programs, although not tested specifically in relation to BLLs, have been effective in improving the cognitive development and social competence of young children, particularly infants from families with low levels of social or economic resources.54 Research demonstrates that children whose development has been delayed or who are at high risk for delay benefit most from interventions applied at an early age.55–57

**Strategies to Prevent and to Reduce Exposure to Lead**

**Major Sources of Exposure**

The CDC and the American Academy of Pediatrics recommend that preventive care for every child should include assessment of environmental history and identification of the occupational lead exposure of household members.2,3,5 The major sources of lead exposure among US children are lead-contaminated dust, deteriorated lead-based paint, and lead-contaminated soil.36,58 Typically, lead contamination of water contributes less to a child’s lead burden than do home and soil sources.58 If additives to water (eg, those used in disinfection pro-
Home-Related Lead Exposure

An estimated 4.1 million homes in the United States (25% of US homes with children <6 years of age) have a lead-based paint hazard. An estimated 68% of US homes built before 1940 have lead hazards, as do 43% of homes built between 1940 and 1959 and 8% of homes built between 1960 and 1977; estimates are higher for homes in the Northeast and Midwest and for homes in which young children reside. Despite considerable attention and resources from federal, state, and local agencies and advocacy groups, publicly available funding has not been able to provide sufficient resources to eliminate all lead paint hazards from US homes. Publicly funded home inspections are most often limited to homes of children with elevated BLLs; the BLL threshold value that prompts an inspection varies according to the state and municipality. Even when a child’s elevated BLL triggers an inspection, public funding for repairs to reduce or to eliminate identified lead hazards typically is not available.

Since 1991, lead hazard-control grant programs through the US Department of Housing and Urban Development Office of Healthy Homes and Lead Hazard Control have provided funding for local and state agencies to reduce lead and other environmental hazards in privately owned, low-income housing. In 2005, the Office of Healthy Homes and Lead Hazard Control allocated $139 million for this purpose, administered through 7 different grant types. Other federal programs provide funding to eliminate lead-based paint hazards in federally assisted housing. The focus of these programs typically is on housing rehabilitation and remediation of lead hazards after children with elevated BLLs are identified, but Department of Housing and Urban Development-funded local programs now include primary prevention interventions that control or eliminate lead before children are exposed.

The CDC is working with the Department of Housing and Urban Development, the US Environmental Protection Agency (EPA), state and local health department lead poisoning prevention grant recipients, and child health and environmental justice advocates to promote primary prevention strategies to reduce exposure to lead. In addition to their traditional role of providing services to children with elevated BLLs, CDC-funded state and local lead poisoning prevention programs have been charged with implementing housing-based primary prevention strategies in their jurisdictions; this involves developing responses to local risks, with a focus on identifying and remediating housing-based lead hazards. The ACCLPP recommendations for essential elements for state and local primary prevention plans have been published previously, as have strategies that have been implemented at the state and local levels to address the problem. As the ACCLPP noted, implementation of state and local primary prevention plans requires (1) targeting of the highest-risk areas, populations, and activities; (2) fostering of political will for jurisdictions to provide adequate levels of funding; (3) expansion of resources for housing remediation and identification and correction of lead hazards; and (4) establishment of a regulatory infrastructure to create and to maintain lead-safe housing and to support the use of lead-safe construction practices. (State prevention plans are available at www.cdc.gov/nceh/lead.)

Certain state and local health departments initiate case management services and home inspections when BLLs reach 10 μg/dL. As more primary prevention strategies are implemented, the number of health departments that pursue home inspections when BLLs reach 10 μg/dL will likely increase. Certain communities have developed online registries to help parents identify homes that are lead-safe or that have lead hazards.

Steps to Identify and Safely Reduce Lead-Based Paint Hazards in Homes

Lead-based paint hazards in homes are important sources of lead exposure. Preventive actions can be implemented to identify and to address these hazards. Tenants can request copies of all lead testing reports for housing sites from landlords at any time. The landlords should have been provided with such information when they purchased the building; compliance with tenant requests for copies of all lead testing reports is required by federal law. In addition, federal regulations require sellers and landlords (1) to disclose the possible presence of lead-based paint in any pre-1978 property and (2) to provide information on known lead-based paint and lead-based paint hazards (eg, by providing the results of any previous evaluations of the property for lead) at the time final agreements are signed for the purchase or rental of most housing built before 1978. Prospective buyers or renters have the opportunity to arrange for a lead inspection or risk assessment by a qualified professional at their own expense; buyers have up to 10 days to check for lead. Furthermore, the law requires sellers, landlords, and renovators to provide buyers, renters, and individuals hiring renovators with an EPA-approved pamphlet (ie, Protect Your Family From Lead in Your Home). To protect their children from lead, parents might choose not to buy or to rent a property or might choose to negotiate remediation of identified lead haz-

Home-Related Lead Exposure

An estimated 4.1 million homes in the United States (25% of US homes with children <6 years of age) have a lead-based paint hazard. An estimated 68% of US homes built before 1940 have lead hazards, as do 43% of homes built between 1940 and 1959 and 8% of homes built between 1960 and 1977; estimates are higher for homes in the Northeast and Midwest and for homes in which young children reside. Despite considerable attention and resources from federal, state, and local agencies and advocacy groups, publicly available funding has not been able to provide sufficient resources to eliminate all lead paint hazards from US homes. Publicly funded home inspections are most often limited to homes of children with elevated BLLs; the BLL threshold value that prompts an inspection varies according to the state and municipality. Even when a child’s elevated BLL triggers an inspection, public funding for repairs to reduce or to eliminate identified lead hazards typically is not available.

Since 1991, lead hazard-control grant programs through the US Department of Housing and Urban Development Office of Healthy Homes and Lead Hazard Control have provided funding for local and state agencies to reduce lead and other environmental hazards in privately owned, low-income housing. In 2005, the Office of Healthy Homes and Lead Hazard Control allocated $139 million for this purpose, administered through 7 different grant types. Other federal programs provide funding to eliminate lead-based paint hazards in federally assisted housing. The focus of these programs typically is on housing rehabilitation and remediation of lead hazards after children with elevated BLLs are identified, but Department of Housing and Urban Development-funded local programs now include primary prevention interventions that control or eliminate lead before children are exposed.

The CDC is working with the Department of Housing and Urban Development, the US Environmental Protection Agency (EPA), state and local health department lead poisoning prevention grant recipients, and child health and environmental justice advocates to promote primary prevention strategies to reduce exposure to lead. In addition to their traditional role of providing services to children with elevated BLLs, CDC-funded state and local lead poisoning prevention programs have been charged with implementing housing-based primary prevention strategies in their jurisdictions; this involves developing responses to local risks, with a focus on identifying and remediating housing-based lead hazards. The ACCLPP recommendations for essential elements for state and local primary prevention plans have been published previously, as have strategies that have been implemented at the state and local levels to address the problem. As the ACCLPP noted, implementation of state and local primary prevention plans requires (1) targeting of the highest-risk areas, populations, and activities; (2) fostering of political will for jurisdictions to provide adequate levels of funding; (3) expansion of resources for housing remediation and identification and correction of lead hazards; and (4) establishment of a regulatory infrastructure to create and to maintain lead-safe housing and to support the use of lead-safe construction practices. (State prevention plans are available at www.cdc.gov/nceh/lead.)

Certain state and local health departments initiate case management services and home inspections when BLLs reach 10 μg/dL. As more primary prevention strategies are implemented, the number of health departments that pursue home inspections when BLLs reach 10 μg/dL will likely increase. Certain communities have developed online registries to help parents identify homes that are lead-safe or that have lead hazards.

Steps to Identify and Safely Reduce Lead-Based Paint Hazards in Homes

Lead-based paint hazards in homes are important sources of lead exposure. Preventive actions can be implemented to identify and to address these hazards. Tenants can request copies of all lead testing reports for housing sites from landlords at any time. The landlords should have been provided with such information when they purchased the building; compliance with tenant requests for copies of all lead testing reports is required by federal law. In addition, federal regulations require sellers and landlords (1) to disclose the possible presence of lead-based paint in any pre-1978 property and (2) to provide information on known lead-based paint and lead-based paint hazards (eg, by providing the results of any previous evaluations of the property for lead) at the time final agreements are signed for the purchase or rental of most housing built before 1978. Prospective buyers or renters have the opportunity to arrange for a lead inspection or risk assessment by a qualified professional at their own expense; buyers have up to 10 days to check for lead. Furthermore, the law requires sellers, landlords, and renovators to provide buyers, renters, and individuals hiring renovators with an EPA-approved pamphlet (ie, Protect Your Family From Lead in Your Home). To protect their children from lead, parents might choose not to buy or to rent a property or might choose to negotiate remediation of identified lead haz-
ards. However, landlords or homeowners might not know whether their property has any lead-based paint or lead hazards.

Lead-based paint hazards are likely to be present in older homes; all homes built before 1978 should be presumed either to have a lead hazard present or to contain intact lead-based paint unless a licensed lead inspector has determined otherwise. Lack of a deteriorated surface decreases the likelihood of lead-contaminated dust being present but does not ensure its absence. Knowledge of the general characteristics of lead-based paint and lead-based paint hazards and their control might help parents to understand their home better (Appendix 1).

Screening for lead dust hazards through dust-wipe testing (ie, standardized collection of dust through wiping of surfaces and measurement of lead collected) can help identify areas of concern. Because lead is not distributed uniformly within a home, wipe testing neither ensures the absence of lead hazards at locations in the home that were not tested nor ensures future protection from lead dust hazards if lead-painted surfaces subsequently deteriorate or are disturbed. Potential sources of future contamination include lead-containing paint on areas disturbed by impact/friction (eg, windows, doors, and floors) and interior migration of lead-contaminated exterior dust and soil. However, identifying lead dust hazards in the home is a first step toward protecting future contamination include lead-containing paint on areas disturbed by impact/friction (eg, windows, doors, and floors) and interior migration of lead-contaminated exterior dust and soil.69 However, identifying lead dust hazards in the home is a first step toward protecting children and might help parents lower lead dust levels in their homes.71 Proper training is recommended for individuals collecting dust wipes, to focus tests on areas at highest risk.62 Parents or property owners who wish to perform dust-wipe sampling may consult their local health or housing departments for advice regarding sampling procedures, interpretation of results, and additional actions based on results.

For a lead-safe environment to be established in older buildings, repair of lead hazards and careful attention to maintenance are necessary. However, local ordinances typically do not require action until a child’s BLL is elevated, and property owners might be unaware of lead hazards or ignore them. Primary prevention is possible only if the focus on safety in older housing is increased and lead hazards are repaired proactively, before a child is exposed. In all pre-1978 properties, owners should use lead-safe work techniques when performing routine maintenance, to decrease the likelihood of lead hazards developing in a home.

Home renovation or repair is known to be a risk factor for increasing or elevated BLLs, principally through exposure to the dust residue generated during the work.34,74–76 All contractors who perform repair and renovation work in older housing should be trained in lead-safe work practices and comply with any state and local requirements governing work with lead paint hazards.77 Property owners performing work themselves should seek expert advice and training to protect themselves and their families.78 Lead-safe work practices include (1) relocating families when the work warrants, (2) minimizing the amount of dust created, (3) containing dust in the work area, (4) cleaning up completely, (5) disposing of waste safely, and (6) performing clearance testing (ie, testing of dust for lead after site cleanup) to ensure that residual lead levels do not exceed EPA standards.80 Families with young children should be restricted from work areas until clearance testing has been performed and the area has been judged to be safe.

In previous evaluation studies, lead dust clearance standards were not low enough to protect children from increased exposure to lead-contaminated dust after lead hazard remediation; as a result, BLLs of children with preremediation BLLs of <25 μg/dL increased after home repairs.81 In 2001, the EPA lead dust clearance standards were lowered to 40 μg/ft² for floors, 250 μg/ft² for window sills, and 400 μg/ft² for window wells.80 No studies have evaluated whether these lower clearance levels protect children with BLLs of <10 μg/dL adequately from ongoing lead exposure. A cross-sectional study estimated that 20% of children with current exposure to floor dust lead at 40 μg/ft² would have BLLs of ≥10 μg/dL.

A study conducted in 1994 to 1999 in 14 US cities, involving 2682 pre-1978 homes, demonstrated reductions in dust lead levels and decreases in children’s BLLs when lead-safe work practices were used during remediation efforts.60,83,84 The study applied lead dust clearance standards substantially less stringent than those currently in place, although clearance floor dust lead levels were generally low (geometric mean: 16 μg/ft²).85 Of the 869 children in that study who were tested within 4 months before home lead remediation and ~7 weeks after remediation, 81 (9.3%) had clinically significant increases (≥5 μg/dL) in BLLs; infants, children of less-educated mothers, and children from homes with greater numbers of preintervention exterior lead hazards were at greatest risk.86 Dust lead levels at clearance were not associated significantly with increases in BLLs. The study listed multiple types of exposures (eg, other homes and parental job exposures) that might have accounted for increased BLLs, but they were not evaluated systematically. Although lead remediation work reduced overall dust lead levels and BLLs, the finding that >9% of children had increases in BLLs of ≥5 μg/dL underscores the need to maintain a high level of vigilance to ensure that children are protected when homes or apartments undergo renovation and repair.

**Educational Strategies**

Lead-exposure–prevention strategies for children with BLLs of <10 μg/dL typically focus on education and promotion of home cleanliness, without further identifying lead hazards or repairing them. Providing low-
income parents with lead-related education via videotape in a pediatric office was demonstrated to be effective in increasing knowledge and parental reports of compliance with lead-prevention actions in the home.97 No studies have evaluated office-based education with accompanying in-home strategies or used children’s BLLs as the outcome measure for an office-based education strategy.

Studies of children at high risk that applied intervention strategies in the home or community demonstrated the failure of education and nonprofessional cleaning conducted alone (ie, in the absence of other measures to reduce lead exposure) in preventing the development of BLLs of \( \geq 10 \) \( \mu \)g/dL. Few studies used prospective designs that included control groups. One study indicated that a highly intensive education program, delivered by community members, that started at birth and lasted for >3 years (28 sessions) decreased the risk of BLLs of \( \geq 10 \) \( \mu \)g/dL by 34%, but this result was not statistically significant.91 Repeated in-home lead-prevention education, even when accompanied by complimentary supplies of cleaning materials, was ineffective in decreasing the incidence of elevated BLLs.92,93 A review of 4 studies94 involving caregiver education93,94 and professional house cleaning95,96 indicated that such low-cost interventions reduced the overall proportions of children with BLLs of \( \geq 15 \) or \( \geq 20 \) \( \mu \)g/dL but the effect on mean BLLs was not statistically significant (\( P > .05 \)).

Intensive cleaning regimens reduce BLLs; in 1 study, biweekly professional cleaning resulted in a 17% decrease in mean BLLs after 1 year.95 However, the benefit of such intense and repeated cleaning was limited to homes without carpets.97 Intensive cleaning can be used without subjecting children to a risk of increased lead exposure resulting from unsafe repair methods (ie, those not in compliance with lead-safe work practices). A single intensive cleaning does reduce levels of lead in dust by 32% to 93%, depending on the surfaces tested and the starting lead concentrations,98 but reaccumulation occurs within 3 to 6 months.99,100

A study that involved children with BLLs of 15 to 19 \( \mu \)g/dL compared the effects of nurse home visits (5 visits in 1 year) accompanied by lead dust tests with those of usual care (1 or 2 visits by an outreach worker in 1 year).71 After 1 year, dust lead levels were significantly lower (\( P < .05 \)) in homes where lead dust tests had been conducted during intervention than in usual-care homes. This finding suggests that dust testing might help parents better understand lead hazards and take action to decrease them. However, changes in dust lead levels were not mirrored by changes in BLLs in this group of children with elevated BLLs.

**BLL-Screening Strategies**

The CDC101 and the American Academy of Pediatrics3 have recommended that health care providers conduct BLL tests for children enrolled in Medicaid and those identified as being at risk on the basis of the state or local screening plan or risk assessment process. Federal policy requires that all children enrolled in Medicaid receive BLL-screening tests at 12 and 24 months of age and that BLL screening be performed for children 36 to 72 months of age who have not been screened previously.102 Despite this, BLL-screening rates for Medicaid-enrolled children have been low (<20%)103 and in certain areas remain ~20%.104 In 1997, the CDC requested that state and local health officials use local community-wide data (eg, BLL prevalence, housing age, and poverty status) to develop plans for BLL screening for their jurisdictions and provide them to clinicians.101 These plans recommend either universal or targeted BLL screening (state and local screening plans are available at www.cdc.gov/nceh/lead/grants/contacts/CLPPP%20Map.htm).

Targeted screening strategies enable clinicians to assess risks for individual children and to recommend BLL testing for the subset of children in the jurisdiction who are thought to be at increased risk for lead exposure. The CDC recommends that risk evaluations be conducted on the basis of factors such as residence in a specific geographic area, membership in a group at high risk, answers on a personal risk assessment questionnaire (which might include local factors such as cultural practices and use of products such as herbal remedies, traditional cosmetics, and imported spices), and other risk factors relevant to the jurisdiction.101

The CDC recommends that locally developed targeted risk assessment and BLL-screening strategies be applied at 1 and 2 years of age.101 Children 36 to 72 months of age who have been identified as being at risk and who have not been screened previously also should receive a BLL test.101 For clinicians in areas that lack a state or local screening plan, the CDC recommends that BLL testing be performed for all children at 1 and 2 years of age and for children 36 to 72 months of age who have not been screened previously.101

Because lead exposure might change with a child’s developmental progress (eg, walking or reaching window sills) or as a result of external factors (eg, family relocation or home remodeling), 2 routine screenings are recommended (at ~1 and ~2 years of age). Among children in Chicago at high risk with BLLs of <10 \( \mu \)g/dL at 1 year of age, 21% had BLLs of \( \geq 10 \) \( \mu \)g/dL when tested again at ~2 years of age.102 That report does not change current CDC recommendations regarding ages for routine BLL testing. However, certain local health departments (eg, those in Chicago, IL; New York, NY; and Philadelphia, PA) recommend BLL screening at younger ages or more frequently.105-107 For example, those departments recommend BLL testing starting at 6 to 9 months of age in high-risk areas, BLL testing at more-frequent intervals (eg, every 6 months) for chil-

---

**PEDIATRICS Volume 120, Number 5, November 2007**

**e1291**
dren <2 years of age, or the provision of additional education and more-rapid follow-up BLL testing for children <12 months of age with BLLs of 6 to 9 μg/dL.

**Personal Lead-Risk-Assessment Questionnaires**

The effectiveness of personal risk assessment questionnaires in identifying children with elevated BLLs has been documented in the scientific literature (Table 2). However, no studies have evaluated the performance of these questionnaires at cutoff levels of <10 μg/dL or their effectiveness in directing counseling or in identifying lead hazards in the home. When used for consecutive samples of patients in clinical settings, the sensitivity of such questionnaires in identifying children with BLLs of ≥10 μg/dL varied considerably according to population. In certain studies, the sensitivity improved if higher cutoff levels were used in the analysis or if the questions used were developed specifically for the population being tested. In general, to identify ~80% of children with BLLs of ≥10 μg/dL, a blood test needed to be performed for more than one half of the children whose risk factors for lead exposure were assessed by using a questionnaire. Multiple studies in populations with low or high prevalence for elevated BLLs concluded that risk assessment questionnaires were not effective in their clinical settings.

**Future Research Needs**

Additional study is needed to assess the effects of BLLs of <10 μg/dL on children. Such research should entail monitoring of large diverse populations, with careful attention to potential confounders and measurements of social factors. Additional research also is needed to evaluate the effectiveness of strategies to decrease exposure to lead. This should include research on the effectiveness of strategies applied in the medical office and the home and interventions provided through medical, public health, and environmental means.

**BLL-screening strategies should be evaluated to determine the most-appropriate ages for screening and the utility of screening strategies applied at the community level.** Evaluations of lead surveillance strategies should test ways to identify changing patterns of environmental risks and subpopulations exposed to established and emerging sources of lead. In addition, better ways to
alert public and clinical health care professionals of changes in exposure sources and patterns and to enhance their responses to such changes through increased surveillance and BLL monitoring of populations identified as being at increased risk for exposure should be identified. Additional studies might provide data that can be used to improve laboratory methods and performance monitoring. This would require the development of criteria to evaluate individual laboratories and mechanisms to provide this information to clinicians.

SUMMARY OF RECOMMENDATIONS

Recommendations for Clinicians

1. Provide anticipatory guidance to parents of all young children regarding sources of lead and help them identify sources of lead in their child’s environment. Obtain an environmental and family occupational history and educate parents about the most common sources of childhood lead exposure for their child and in their community. Encourage parents to identify lead hazards and sources in their homes and to reduce their child’s potential for exposure to lead, including the safe implementation of control measures before BLLs increase. Warn parents about the dangers posed by unsafe renovation methods and urge them to be cognizant of the possibility of new and reemerging sources of lead in children’s environments. Direct parents to local, state, and federal agencies and organizations for information, particularly concerning methods to identify lead hazards and to repair them safely (Appendix 2).

2. Help parents to understand the uncertainty of BLL values and potential reasons for their fluctuation, including errors introduced by the sampling methods and laboratory-, age-, and season-related exposures.

3. Assess all children for developmental and behavioral status and seek additional evaluation and therapy to reduce developmental or behavioral problems, as necessary. Consider the potential influences of lead when conducting developmental screening. For children with multiple developmental risk factors that might include lead exposure, consider more-frequent developmental surveillance or conduct more-extensive developmental evaluations.

4. Discuss with parents the potential impact of lead on child development and promote strategies that foster optimal development, including encouraging parents to influence their child’s development positively by providing nurturing and enriching experiences. For all children from families with low levels of economic and social resources who are living in areas where exposure to lead is likely, promote participation in early enrichment programs regardless of the child’s BLL.

5. Whenever possible, use laboratories that can achieve routine performance of ±2 μg/dL for BLL analysis. Evaluate laboratory performance by reviewing the laboratory’s quality control chart or statistical quality control summary.

6. Review office procedures and policies to ensure that lead-exposure risk assessment or BLL screening is performed for all children as required by state or local health officials or as recommended by the CDC. Consider the child’s age, season of testing, and exposure history when deciding when to obtain follow-up BLL tests. For a child whose BLL is approaching 10 μg/dL, more-frequent BLL screening (ie, more often than annually) might be appropriate, particularly if the child is <2 years of age, was tested at the start of warm weather (when BLLs tend to increase), or is at high risk for lead exposure.

7. Perform a diagnostic BLL test for all children suspected of having lead exposure or an elevated BLL and institute the recommended management guidelines if a child’s BLL increases to ≥10 μg/dL.

8. Become informed about lead-exposure–prevention strategies of local or state health departments and partner with public health agencies, community groups, and parents to work toward establishing lead-safe environments in homes and schools for all children and reducing exposure to lead from all sources. Advocate for the expansion of services that foster primary lead poisoning prevention.

Recommendations for Government Agencies

1. Increase efforts to resolve lead-based paint hazards safely before children are exposed.

2. Expand services that promote primary lead poisoning prevention and develop systems that enable clinicians and parents to learn about such services.

3. Develop and implement strategies to encourage the safe elimination of lead hazards in properties, using trained workers and lead-safe work practices, in compliance with federal, state, and local regulations.

4. Establish jurisdictional policies that mandate ensuring lead safety in housing and enforce these mandates.

5. Develop and apply systematic approaches to prevent exposure to even small amounts of lead in food or consumer products, particularly when safer alternatives are available.

6. Promote implementation of state and local primary prevention plans that target areas, populations, and activities of highest risk; expand resources for hous-
ing remediation; identify and correct lead hazards; and establish a regulatory infrastructure to create and to maintain lead-safe housing and to support the use of lead-safe construction practices.

7. Expand the availability of and promote the use of early enrichment programs for all children from families with low levels of economic and social resources who are living in areas where exposure to lead is likely.

8. Promote and fund additional research to evaluate the effects of lead at BLLs of <10 μg/dL and to evaluate strategies to identify and to reduce exposure or the potential for exposure to lead, including strategies applied in medical offices and in homes.

APPENDIX 1. TIPS FOR REDUCING LEAD-BASED PAINT AND LEAD-BASED PAINT HAZARDS

- The concentration of lead is generally highest in lead-based paint on exterior surfaces.
- Among interior surfaces, windows are most likely to have the highest lead content.
- Interior surfaces can become contaminated from exterior sources or common areas.
- Lead-based paint on impact/friction surfaces (eg, windows, doors, and floors) deteriorates as paint is disturbed during use.
- Lack of a deteriorated surface does not ensure the absence of lead-contaminated dust, although it decreases the risk.
- Renovation, remodeling, and repainting can increase dust lead levels significantly.
- Vacuum methods (using a traditional vacuum or a high-energy particulate air-filtered vacuum) do not decrease lead levels on soiled carpets or upholstery enough to achieve safe levels.
- Creating smooth cleanable surfaces helps achieve lower dust lead levels.
- Treatments addressing lead-contaminated exterior dust/soil and building exterior lead hazards can contribute to lower dust lead levels in entryway and home interior locations.
- Safely addressing interior, exterior, and soil lead hazards in an integrated manner is most beneficial in establishing lasting, lead-safe environments.

ACKNOWLEDGMENTS

The ACCLPP membership included chairperson Carla Campbell, MD, (Children’s Hospital of Philadelphia, Philadelphia, PA); executive secretary Mary Jean Brown, ScD, RN (Division of Environmental and Emergency Health Services, National Center for Environmental Health/Agency for Toxic Substances and Disease Reg-

APPENDIX 2 Guide to Resources for Parents

<table>
<thead>
<tr>
<th>Agency/Organization</th>
<th>Specific Resources</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Lead Information Center</td>
<td>Multiple general publications</td>
<td>Telephone: 800–424–LEAD (800–424–5323); fax: 585–232–3111; Web site: <a href="http://www.epa.gov/lead/pubs/nlic.htm">www.epa.gov/lead/pubs/nlic.htm</a>; e-mail: see Web site</td>
</tr>
<tr>
<td>CDC, Lead Poisoning Prevention Branch</td>
<td>Questions and answers; fact sheets; lead screening plans by state and area; links to other sources</td>
<td>Telephone: 770–488–3300; fax: 770–488–3635; Web site: <a href="http://www.cdc.gov/nceh/lead">www.cdc.gov/nceh/lead</a>; e-mail: <a href="mailto:leadinfo@cdc.gov">leadinfo@cdc.gov</a></td>
</tr>
<tr>
<td>EPA, Office of Pollution Prevention and Toxics</td>
<td>Lead in Your Home: A Parent’s Reference Guide; Reducing Lead Hazards When Remodeling Your Home; Protect Your Family From Lead in Your Home; Is There Lead in My Drinking Water?; Lead Poisoning and Your Children; other materials</td>
<td>Telephone: 202–566–0500; fax: 202–566–0469; Web site: <a href="http://www.epa.gov/lead">www.epa.gov/lead</a>; e-mail: see Web site</td>
</tr>
<tr>
<td>Alliance for Healthy Homes</td>
<td>Lead Safety Tips for Tenants; Lead-Safe Painting and Renovation; links to other sources</td>
<td>Telephone: 202–543–1147; fax: 202–543–4466; Web site: <a href="http://www.afhh.org">www.afhh.org</a>; e-mail: <a href="mailto:afhh@afhh.org">afhh@afhh.org</a></td>
</tr>
<tr>
<td>National Center for Healthy Housing</td>
<td>A Guide to Working Safety With Residential Paint; links to other sources</td>
<td>Telephone: 410–992–0712; fax: 410–715–2310; Web site: <a href="http://www.centerforhealthyhousing.org">www.centerforhealthyhousing.org</a>; e-mail: <a href="mailto:nchh@enterprisefoundation.org">nchh@enterprisefoundation.org</a></td>
</tr>
<tr>
<td>National Library of Medicine, National Institutes of Health</td>
<td>Links to other sources, including Spanish-language materials</td>
<td>Web site: <a href="http://sis.nlm.nih.gov/enviro/lead.html">http://sis.nlm.nih.gov/enviro/lead.html</a></td>
</tr>
</tbody>
</table>
istry, CDC, Atlanta, GA); members Magaly C. Angeloni, MBA (Rhode Island Department of Public Health, Providence, RI), Valerie Charlton, MD (California Department of Health, Richmond, CA), Walter S. Handy, Jr, PhD (Cincinnati Health Department, Cincinnati, OH), Ing Kang Ho, PhD (University of Mississippi Medical Center, Jackson, MS), Valarie Johnson (Urban Parent to Parent, Rochester, NY), Linda Kite, MBA (Healthy Homes Collaborative, Los Angeles, CA), Jessica Leighton, PhD (New York City Department of Health and Mental Hygiene, New York, NY), George G. Rhoads, MD (University of Medicine and Dentistry of New Jersey, Piscataway, NJ), Catherine M. Slota-Varma, MD (deceased; Medical College of Wisconsin, Milwaukee, WI), Wayne R. Snodgrass, MD, PhD (University of Texas Medical Branch, Galveston, TX), Kevin U. Stephens, Sr, MD, JD (New Orleans Department of Health, New Orleans, LA), Helen J. Binns, MD (member in 2002–2004; Feinberg School of Medicine, Northwestern University, Chicago, IL), Kimberly M. Thompson, ScD (member in 2002–2005; Harvard University, Boston, MA); nonvoting federal members Phyllis Stubbs-Wynn, MD (Maternal and Child Health Bureau, Health Resources and Services Administration, Washington, DC), Michael Bolger, PhD (US Food and Drug Administration, Washington, DC), David Jacobs, PhD (member in 1996–2004; US Department of Housing and Urban Development, Washington, DC), Warren Friedman, PhD (US Department of Housing and Urban Development, Washington, DC), Jacqueline E. Mosby, MPH (EPA, Washington, DC), Walter Rogan, MD (National Institute of Environmental Health Sciences, Washington, DC), Robert J. Roscoe, MS (National Institute for Occupational Safety and Health, CDC, Cincinnati, OH), Lori E. Saltzman, MS (US Consumer Product Safety Commission, Washington, DC); and nonvoting liaison representatives Steve M. Hays (American Industrial Hygiene Association, Nashville, TN), Ezatollah Keyvan-Larijani, MD, DrPH (Council of State and Territorial Epidemiologists, Baltimore, MD), Pat McLaine, MPH (representative in 1998–2005; National Center for Healthy Housing, Columbia, MD), Jonathan Wilson, MPP (National Center for Healthy Housing, Columbia, MD), Benjamin Gitterman, MD (American Public Health Association, Washington, DC), Routt Reigart II, MD (representative in 1997–2004; American Academy of Pediatrics, Charleston, SC), George C. Rodgers, Jr, MD, PhD (American Association of Poison Control Centers, Georgetown, IN), Jan Towers, PhD (American Academy of Nurse Practitioners, Gettysburg, PA), Anne M. Wengrovitz, MPH (Alliance for Healthy Homes, Washington, DC), and Calvin B. Johnson, MD (American State and Territorial Health Officials, Harrisburg, PA).

Helpful suggestions were provided by Patrick J. Parsons, PhD (Lead Poisoning/Trace Elements Laboratory, Wadsworth Center, New York State Department of Health, Albany, NY).

REFERENCES

18. Parsons PJ, Geraghty C, Verostek MF. An assessment of contemporary atomic spectroscopic techniques for the dete-
nation of lead in blood and urine matrices. Spectrochim Acta B. 2001;56:1593–1604


32. Chen A, Dietrich KN, Ware JH, Radcliffe J, Rogan WJ. IQ and blood lead from 2 to 7 years of age: are the effects in older children the residual of high blood lead concentrations in 2-year-olds? Environ Health Perspect. 2005;113:597–601


39. Lanphear BP, Matte TD, Rogers J, et al. The contribution of lead-contaminated house dust and residential soil to chil-


66. Lead: requirements for disclosure of known lead-based paint and/or lead-based paint hazards in housing. 42 USC §4852d (1992)


77. US Environmental Protection Agency. *EPA Model Renovation Training Course Minimizing Lead-Based Paint Hazards During Renovation, Remodeling, and Painting.* Washington, DC: US Environmental Protection Agency; 2000. Publication EPA 747-B-00-005/6


91. Jordan DM, Yuse BL, Robinson LL, Hannan P, Deinard AS. A randomized trial of education to prevent lead burden in chil-


Guidelines for Adolescent Depression in Primary Care (GLAD-PC): I. Identification, Assessment, and Initial Management

Rachel A. Zuckerbrot, MD, Amy H. Cheung, MD, Peter S. Jensen, MD, Ruth E. K. Stein, MD, Danielle Laraque, MD, and the GLAD-PC Steering Group

Financial Disclosure: Dr Cheung is on the speakers’ bureau of Eli Lilly; Dr Jensen has received several unrestricted educational grants from Eli Lilly, McNeil, and Janssen-Ortho, is a consultant for Shire-Richwood, UCB Pharma, McNeil, and Janssen-Ortho, and is on the speakers’ bureau for UCB Pharma, McNeil, and Janssen-Ortho. The other authors have indicated they have no financial relationships relevant to this article to disclose.

Endorsements: The Canadian Paediatric Society, the Society for Adolescent Medicine, the Canadian Association for Adolescent Health, the National Association for Pediatric Nurse Practitioners, the Society for Developmental and Behavioral Pediatrics, the American Academy of Child and Adolescent Psychiatry, the Canadian Academy of Child Psychiatry, the Canadian Psychiatric Association, the College of Family Medicine of Canada, the National Alliance on Mental Illness, the Mental Health Association of New York City, the National Mental Health Association (now known as Mental Health America), the Depression and Bipolar Support Alliance, and the Federation of Families for Children’s Mental Health have endorsed these guidelines. Endorsements from the American Academy of Pediatrics and the Canadian Psychological Association are pending. The American Academy of Family Physicians, the American Medical Association, and the American Psychological Association have been involved in the development of the guidelines but do not endorse external guidelines.

ABSTRACT

OBJECTIVES. To develop clinical practice guidelines to assist primary care clinicians in the management of adolescent depression. This first part of the guidelines addresses identification, assessment, and initial management of adolescent depression in primary care settings.

METHODS. By using a combination of evidence- and consensus-based methodologies, guidelines were developed by an expert steering committee in 5 phases, as informed by (1) current scientific evidence (published and unpublished), (2) a series of focus groups, (3) a formal survey, (4) an expert consensus workshop, and (5) draft revision and iteration among members of the steering committee.

RESULTS. Guidelines were developed for youth aged 10 to 21 years and correspond to initial phases of adolescent depression management in primary care, including identification of at-risk youth, assessment and diagnosis, and initial management. The strength of each recommendation and its evidence base are summarized. The identification, assessment, and initial management section of the guidelines includes recommendations for (1) identification of depression in youth at high risk, (2) systematic assessment procedures using reliable depression scales, patient and caregiver interviews, and Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition criteria, (3) patient and family psychoeducation, (4) establishing relevant links in the community, and (5) the establishment of a safety plan.

CONCLUSIONS. This part of the guidelines is intended to assist primary care clinicians in the identification and initial management of depressed adolescents in an era of great clinical need and a shortage of mental health specialists but cannot replace clinical judgment; these guidelines are not meant to be the sole source of guidance for adolescent depression management. Additional research that addresses the identification and initial management of depressed youth in primary care is needed, including empirical testing of these guidelines.
MAJOR DEPRESSION IN adolescents is recognized as a serious psychiatric illness with extensive acute and chronic morbidity and mortality.\textsuperscript{1,2} Research shows that only 50% of adolescents with depression are diagnosed before they reach adulthood.\textsuperscript{3} In primary care (PC), as many as 2 in 3 depressed youth are not identified by their PC clinicians and do not receive any kind of care.\textsuperscript{4,5} Even when diagnosed by PC physicians, only half of these patients are treated appropriately.\textsuperscript{3} Furthermore, rates of completion of specialty mental health referral for youth with a recognized emotional disorder from general medical settings are quite low (J. V. Campo, MD, written communication, 2006).

In view of the shortage of mental health clinicians and barriers to children having access to mental health professionals, the well-documented need for PC clinicians to learn how to manage this condition, the increasing evidence base available to guide clinical practice, increased selective serotonin reuptake inhibitor–prescribing rates in pediatric PC,\textsuperscript{6,7} and new evidence that a multifaceted approach with mental health consultation may improve the management of depression in PC settings,\textsuperscript{8–11} guidelines may be a necessary first step in the identification and management of depression in adolescents in PC. Unfortunately, no depression-management guidelines have been developed for use in the PC setting in the United States or Canada.

Although additional randomized, controlled, clinical trial (RCT) information is urgently needed to guide PC clinicians in optimal management approaches, such studies often take years to complete, and many critical PC adolescent depression-management questions have not been, and will likely never be, addressed in completed or ongoing studies. To address this gap and meet the needs of PC clinicians and families who are on the “front lines” with few mental health resources available, this report and its companion article\textsuperscript{12} constitute the first-ever evidence- and expert consensus–derived guidelines to guide PC clinicians’ management of adolescent depression. These guidelines are also accompanied by a toolkit (available at no cost for download at www.glad-pc.org; see Appendix).

Over the last 3 years, the Center for the Advancement of Children’s Mental Health at Columbia University and the Sunnybrook Health Sciences Center at the University of Toronto joined forces with the New York Forum for Child Health, New York District II and New York Chapters 1 through 3 of the American Academy of Pediatrics and, more recently, the REACH Institute, along with leading experts across the United States and Canada, to address the need for a synthesis of knowledge in this area. The result of this initiative was the development of the Guidelines for the Management of Adolescent Depression in Primary Care (GLAD-PC). These guidelines are based on available research and consensus of experts in depression and in PC. In this article, we present the summary result of literature reviews of the available data and the recommendations on the identification and assessment of depression in PC settings; in our accompanying report,\textsuperscript{12} we present the results of the reviews and recommendations on treatment (psychotherapy, psychopharmacology, and pediatric counseling) and ongoing management. Although very few studies have addressed adolescent depression identification and management in PC settings, many PC clinicians are already attempting to change their clinical practices; thus, a great need exists to develop and disseminate methods and tools for assisting PC clinicians in managing adolescent depression.

Major depressive disorder (MDD) is a specific diagnosis described in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV),\textsuperscript{13} which includes symptoms of low mood, anhedonia, and other neurovegetative symptoms (ie, insomnia, decreased concentration, low energy, etc). Other types of depression exist, including dysthymia, subthreshold forms, or those that occur as part of bipolar disorder or other mental illness. Although the evidence for the psychopharmacology recommendations in the accompanying article\textsuperscript{12} focus exclusively on MDD, the recommendations around identification, assessment, and initial management can be applied to other forms of depression as well.

Our guidelines also distinguish between mild, moderate, and severe forms of MDD. The DSM-IV depression criteria include 9 specific symptoms that have been shown to cluster together, run in families, and have a genetic basis;\textsuperscript{14–18} and a large body of evidence accumulated over time now supports the internal consistency of depressive symptoms and the validity of the major depression construct, based on the validational criteria for all psychiatric diagnoses.\textsuperscript{14} According to the DSM-IV, severity of depressive disorders can be based on symptom count. This commonly used method to define depression severity has been used in large population-based studies\textsuperscript{19} and may be particularly relevant in PC settings in which less-severe clinical presentations of depression may be more common. Thus, mild depression may be characterized on the basis of lower scores in standardized depression scales with shorter duration of symptoms or meeting minimal criteria for depression. Following the DSM-IV, mild depression might be defined as 5 to 6 symptoms that are mild in severity. Furthermore, the patient might experience only mild impairment in functioning.

In contrast, depression might be deemed to be severe if a patient experiences all of the depressive symptoms listed in the DSM-IV. Depression might also be considered severe if the patient experiences severe impairment in functioning. Moderate depression falls between these 2 categories. In general, however, even if not all 9 DSM-IV–defined symptoms of depression are present, for the purposes of these guidelines, an adolescent who meets at
least 5 criteria for the diagnosis of MDD should be considered to be in the severe category if he or she presents with a specific suicide plan, clear intent, or recent attempt; psychotic symptoms; or severe impairment in functioning (such as being unable to leave home).

These guidelines were developed for PC clinicians who are in a position to identify and assist depressed youth in their practice settings. Although the age range of 10 to 21 years may encompass preteens, adolescents, and young adults in specific instances, this age range was chosen to include those who might be developmentally “adolescent.” Research that supports adult depression guidelines includes adults 18 years and older. Much of the adolescent depression research focuses on children 18 years and younger. However, because adolescent medicine clinicians and school health clinicians often see patients until they are 21 years old, we have included the older adolescents. However, a primary caregiver faced with an adolescent between the ages of 18 and 21 can choose to use either adult or adolescent depression guidelines on the basis of the developmental status of the adolescent and their own comfort and familiarity with each set of guidelines.

**METHODS**

The following recommendations were developed on the basis of a synthesis of expert consensus- and evidence-based research-review methodologies. We compiled the necessary information to develop these recommendations in 5 phases.

1. To understand the problems and obstacles faced by PC clinicians regarding the management of adolescent depression, we first conducted focus groups with PC clinicians and youth patients and their family members to review issues pertinent to the PC management of depression.

2. Systematic literature reviews were conducted in each of 5 key areas in which recommendations were subsequently developed. Whenever possible, these reviews focused on identifying empirical evidence that was developed within child/adolescent PC settings. When PC studies were unavailable, research from specialty mental health care was reviewed. In all 5 review instances, the GLAD-PC Steering Group first determined the existence of all high-quality, previously published, systematic evidence-based reviews that met the following criteria: (a) clear definition of search terms from Medline, including words and word roots; (b) explicit delineation of years searched; (c) exclusion of non–English-language studies; (d) physical review and reading of search-identified titles and abstracts; and (e) selection, review, and reading of possibly relevant articles before determination of final inclusion. When more than 1 systematic evidence-based review was identified for a given area, all reviews were drawn on to identify relevant articles for potential inclusion. More than 1 systematic evidence-based review was available for the areas of efficacy of psychotherapeutic interventions for youth MDD and efficacy of pharmacologic treatments for youth MDD. For all reviews, when appropriate, we updated the review for any ensuing years transpired since the latest review by using these same 5 methods. When systematic reviews were not available for a given area, the GLAD-PC team conducted a systematic review by using Medline (from inception to 2004/2005) and the criteria described above. Reviews were guided by members of the GLAD-PC Steering Committee, which comprised leading experts in each of these areas.

To address the first key area regarding the identification and assessment of adolescent depression in PC, a systematic evidence review was conducted to identify all available evidence about adolescent depression identification in PC, as well as information regarding current practices. This review has since been published. Because of limited information about depression assessment and screening measures in PC specifically, we also reviewed adolescent-screening instruments/tools previously used in psychiatric or community populations. Beginning from 2 previous systematic evidence reviews, the GLAD-PC team performed an additional systematic review from 1998 to 2004.

To address the second key area regarding the initial management of adolescent depression in PC, a systematic evidence review was conducted to identify all available evidence about interventions for adolescent depression in PC and has since been published as well. Other evidence for the initial management of adolescent depression in PC came from systematic evidence reviews that addressed the chronic illness model, systems of care, and safety planning for suicidal patients.

3. On the basis of the questions and issues identified during the focus groups and the literature reviews, we developed a survey to answer questions regarding critical issues in PC management of adolescent depression that have not been answered in the empirical literature. The survey questions were developed and reviewed by clinical and research experts in the area of mental health and PC. Using this survey, research and clinical experts were surveyed on their depression assessment and management recommendations. Depression clinical/research experts (N = 81) from Canada and the United States were asked to complete the 34-item study survey. Of these items, 3 questions dealt with the identification and diagnosis of depression. Subjects were chosen by using 1 of 4 criteria: (a) membership in child and adolescent psychiatric organizations in Canada and the United States including their academies of child and adoles-
cent psychiatry; (b) recipient of federal grants for related research; (c) lead author of at least 2 articles on clinical research in the area from 1999 to 2004 on the basis of Medline citations; or (d) key PC clinical and research leader with expertise in the area of guideline development and/or emotional and behavioral disorders that present in PC settings. Complete survey results \((n = 76)\) will be presented in a subsequent peer-reviewed article and are available from the authors on request.

4. An expert consensus workshop was held in July 2004 with 81 North American experts on depression, clinical pediatrics, quality improvement, mental health policy, and health economics. Published data from the literature review, unpublished high-quality research currently in process of publication, and the results of the survey were presented to guide the initial discussion and consensus process.

5. Guidelines were developed on the basis of multiple iterations shared among a small group of core writers, guidance of the larger steering committee, and ultimate input of all consensus-conference attendees to obtain full ownership of the final product. The results of this process are presented below.

On the basis of the 5-step method, 2 guidelines were developed to address different areas of adolescent depression management in PC settings: (1) identification, assessment, and initial management and (2) treatment and ongoing management. This part of the guidelines focuses on identification, assessment, and initial management. Each section of both guidelines is composed of individual recommendations followed by a brief rationale that refers to available empirical findings and experts’ consensus opinion on which the recommendations were based. Each recommendation is graded on the basis of the Oxford Centre for Evidence-Based Medicine grade of evidence (A–D) system (see www.cebm.net/levels-of-evidence.asp). In addition, the strength of each recommendation, in terms of the extent to which experts agreed that the recommendation is highly appropriate and a “first-line” practice, was reached for each recommendation. Recommendation strength was rated in 4 categories: very strong (>90% agreement), strong (>70% agreement), fair (>50% agreement), and weak (<50% agreement). The recommendations in the guidelines were developed only in areas of management that had at least “strong agreement” among experts.

RESULTS

Literature Reviews: Identification and Assessment

Twenty-five articles were located that discussed specific identification methods for adolescent depression in pediatric PC. Only 10 of these articles presented psychometric data of any kind, such as sensitivity, specificity, positive predictive value, negative predictive value, and area under the curve (a full table is available in the published review). One of the 25, plus an additional 5 survey articles, dealt with current identification practices by PC physicians. According to those data, most PC clinicians rely on the use of presenting complaints and family concerns to identify depressed youth. Likewise, other surveys confirmed that very few pediatric providers have instituted a systematic assessment or case-finding tool to identify adolescent depression.

Despite clinicians’ principal reliance on adolescent and parental chief complaints and physician interview in current practice, the authors of the review found that the use of these methods alone underidentify adolescent depression. Even physicians who are trained in the use of mnemonics to guide interviews underidentify adolescent depression. Instead, only by asking patients directly about depression and suicide (versus relying on them to volunteer the information) does one reliably improve case finding and the psychometric quality of diagnostic data. Using systematic assessment methods with depression-specific questions seems to provide the best identification results.

Many steps are involved in implementing systematic depression-identification procedures within a busy practice setting (eg, training office staff in the use of procedures, adjusting other paperwork demands to fit the depression-identification procedures, ensuring that providers review the information, teaching providers how to use the information, and determining whether the procedures actually benefit youth). Although no study has documented the feasibility and outcomes of taking all of these steps within a single study, evidence from the review cited above as well as a recently published pilot study suggests that each of these steps can be implemented in real-world settings (ie, training office staff, ensuring PC providers review any results, and establishing that appropriate counseling and/or needed mental health referrals are made). Finally, and perhaps most importantly, although no study compared outcomes between screened and not-screened groups, 1 study did demonstrate that an identification program in PC when combined with high-quality depression treatment actually yields better outcomes than treatment-as-usual conditions (when no high-quality depression treatment is available).

Because of limited information about depression assessment and screening measures in PC specifically, we also reviewed adolescent-screening instruments/tools previously used in psychiatric or community populations. Beginning from 2 previous systematic evidence reviews, the GLAD-PC team performed an additional systematic review for the years 1998–2004 and found 15 studies with psychometric performance data on depression instruments in English-speaking adolescents. A table that presents these 15 studies, the gold-standard
diagnostic assessment used, sensitivity, specificity, positive predictive value, negative predictive value, prevalence, and study population is available on request. Commonly used adolescent-screening instruments included the Beck Depression Inventory, the Reynolds Adolescent Depression Screen, and the Mood and Feelings Questionnaire. Sensitivities of these more common instruments ranged from .70 to .90, and specificities ranged from .39 to .90. In addition, the Kutcher Adolescent Depression Scale, a newer 6-item instrument, had sensitivity and specificity of .81 and .86, respectively, in a school population with a positive predictive value of .39 and a negative predictive value of .98 given a 10% prevalence.

In summary, no perfect depression screening/assessment tool exists, but a number of adolescent depression assessment instruments do possess adequate psychometric properties to commend their use in depression detection and assessment. Thus, it is reasonable to expect that depression detection in PC can be improved by the use of self- or parent-report checklists. Reliance on adolescent self-report depression checklists alone will lead to substantial numbers of false-positive and false-negative cases. Instead, optimal diagnostic procedures should combine the use of depression-specific screening tools as diagnostic aids buttressed by follow-up clinical interviews in which one obtains information from other informants (eg, parents) and reconciles discrepant information to arrive at an accurate diagnosis and impairment assessment before treatment. For more information about rating scales and cutoff scores, please refer to the GLAD-PC toolkit (available at www.glad-pc.org; see Appendix).

Initial Management of Adolescent Depression
A GLAD-PC review of adolescent depression interventions performed in PC found 4 articles that directly discussed interventions for adolescent depression in PC, only 1 of which reported an intervention by a PC staff member as opposed to a mental health worker. This identified RCT evaluated the effectiveness of a PC-delivered consultation intervention that invited teens from 8 general practices in Britain to discuss their health concerns with a PC nurse who provided individual consultations. In addition to discussing adolescents’ general health concerns, the PC nurses offered mental health referrals when they deemed it appropriate. Posthoc analyses indicated that among teens with high Center for Epidemiological Studies Depression Scale for Children (CES-DC) scores, those who were randomly assigned to the PC nurse consultation had lower CES-DC scores on follow-up than adolescents with high CES-DC scores who were not randomly assigned to the consultation, which suggests that a PC-delivered intervention may be useful in addressing depression symptoms and/or similar emotional issues (because these patients were not diagnosed with a depressive disorder per se), perhaps by the consultation itself or by the resulting referral to mental health.

Although the study discussed above was the only one that dealt with an intervention by a PC provider for adolescent depression, the large multisite trial conducted by Asarnow et al showed that improving links between PC and mental health will result in better outcomes for depressed adolescents who present in PC.

In addition, on behalf of the GLAD-PC team, Stein et al reviewed the literature on psychosocial interventions for anticipatory guidance. No RCTs or evidence-based reviews were found. Citing earlier literature reviews in the area of injury prevention and anticipatory guidance; Stein et al found some limited evidence that anticipatory guidance strategies such as education and counseling in the PC setting can be effective.

Another area reviewed by Stein et al involved psychosocial interventions for improved adherence. An evidence review by Lemanek et al on asthma adherence suggested that some educational and behavioral strategies are “probably efficacious” in creating change. In addition, a study that used cognitive behavioral strategies suggested that diabetic adherence can also be improved.

In addition, our team searched the Cochrane database for systematic reviews of all types of interventions implemented in the adherence arena. These reviews suggested that only complex, multifaceted approaches that included convenient care, patient education, reminders, reinforcement, counseling, and additional supervision by a member of the care team were effective in improving adherence in different chronic medical conditions including asthma, hypertension, diabetes, and adult depression. These complex models have also moved beyond the somewhat paternalistic traditional model to a model of shared decision-making between the provider and patient. However, improvements in treatment outcomes remain small even with these complex and resource-intensive interventions. In the pediatric literature, research regarding adherence commonly involved interventions that targeted both patients and their families. Several key components have been identified that may improve compliance/adherence, including patient self-management/monitoring, patient/family education/support, and the setting and documentation of achievement of management goals. The identification and periodic review of short- and long-term goals provides an individualized plan that both the provider and the patient and family can follow over time. However, current research evidence suggests that more complex interventions that include shared decision-making between provider and patient are likely to have the greatest impact on both adherence and treatment outcomes. This kind of coordinated care has been designated the PC medical home and is discussed further in this section.
Factors that pertain to the linkages between/among organizations have been studied as well. For example, the concept of “system of care” was first described in 1986 in a monograph by Stroul and Friedman and used to support the Children and Adolescent Service System Program; their monograph outlined the ideal model for an integrated system of care for children and youth with severe emotional disturbances. Over the last 2 decades, numerous examples of comprehensive, coordinated, community-based systems of care have been implemented and evaluated. Most of them showed no improvement in patient outcomes but significant improvement in other areas such as patient/family satisfaction, decreased wait time, and more appropriate care.

In the chronic care medical literature, PC professionals have been encouraged to provide a medical home for their patients with chronic conditions. In effect, professionals are expected to remain accessible to patients and families through periods of quiescence and medical crises, coordinate care with other health care professionals, advocate with third-party payers, and provide continuity of care to prevent long-term consequences of chronic illnesses. A review of the evidence for the medical home was conducted by Cooley, who found that although no studies of the outcomes of the broad application of the medical home exist, some evidence exists for positive outcomes for different aspects of the concept. Again, no RCTs exist that we could find. In 1 study, children without a medical home were twice as likely to delay or forego needed care and to have unmet health care service needs. Two other reviews focused on the impact of coordination of care and continuity of care. As a recent vaccination study demonstrated, just having a medical home available to a patient does not necessarily mean that the patient will make full use of the available services. Thus, as outlined in a recent American Academy of Pediatrics policy statement, the medical home must provide care coordination and help patients to make use of available resources.

Although the “system of care” focuses on the overall coordination of care involving many agencies that are involved in the care of youth (child welfare services, schools, etc), a crucial link in that system and for adolescent depression is that between mental health and PC. Currently, no specific literature addresses the issue of referral to specialty care of adolescent patients with depression (such as which subgroups of patients would benefit from referral to mental health professionals). Studies have been conducted with adults with depression that demonstrate that increased collaborative care between mental health and PC professionals is needed to improve the care of patients with depression in PC. The adult literature shows the importance of a close working relationship between mental health specialists and PC clinicians in the PC setting. Different models of collaboration have been shown to be effective in the adult depression literature, including the use of case managers in PC practices and consultation by on-site mental health clinicians. Other models include shared care and telephone consultation on an ad hoc basis. Although these models suggest improved outcomes for both clinicians and patients, there are significant barriers to the successful implementation of these collaborative models, including funding deficiencies and shortage of mental health providers. Financial and other incentives for both PC and mental health clinicians to develop these models and obtain the training needed to function within these models are essential. Increased development of skills needed for collaborative care and training in mental health may also be addressed at earlier stages of training such as during residency for both PC and mental health clinicians.

Safety Planning

Safety planning with depressed suicidal or potentially suicidal adolescent patients usually consists of instructing the family to remove lethal means, instructing the family to monitor for risk factors for suicide, engaging the potentially suicidal adolescent in his or her treatment, providing adolescents with mutually agreeable and available emergency contacts should they find themselves with increasing suicidality, and establishing clear follow-up. Our review of the literature found no trials that have studied the impact of or how to conduct any of these aspects of safety planning with depressed adolescents. No studies have examined the benefits or risks of a safety contract. Some studies have suggested that limiting access to firearms or other lethal means can decrease suicide by those methods, but the evidence is still unclear as to whether, on a broader population level, restricting access to certain lethal methods results in an overall decrease of suicide rates. In addition, a study by Brent et al found that families of depressed adolescents are frequently noncompliant with recommendations to remove firearms from the house. Yet, a small prospective follow-up of patients seen in an emergency department for mental health concerns found that the majority of families removed or secured lethal means (firearms, alcohol, prescription medications, and over-the-counter medications) after injury-prevention education in the emergency department, whereas no families who did not receive injury-prevention education did so. Some limited evidence suggests that quick and consistent follow-up with a team approach will be most helpful in increasing compliance and engagement among suicidal patients.

GUIDELINES

Identification

Recommendation 1: Patients with depression risk factors (such as history of previous episodes, family history, other psychiatric
disorders, substance abuse, trauma, psychosocial adversity, etc) should be identified (grade of evidence: C; strength of recommendation: very strong) and systematically monitored over time for the development of a depressive disorder (grade of evidence: C; strength of recommendation: very strong).

Although most PC clinicians believe it is their responsibility to identify depression in their adolescent patients, evidence suggests that only a fraction of these youth are identified when they present in PC settings, and only 50% of depressed adolescents are diagnosed before reaching adulthood.323 As part of overall health care, PC clinicians should routinely monitor the psychosocial functioning of all youth, because problems in psychosocial functioning may be an early indication of a variety of problems, including depression. For those at known increased risk for depression, PC clinicians should use systematic, effective identification strategies. Risk factors that clinicians may use to identify those who are at high risk for depression include a personal history and/or family history of (1) depression, (2) bipolar disorder, (3) suicide-related behaviors, (4) substance abuse, and (5) other psychiatric illness, or (6) significant psychosocial stressors such as family crises, physical and sexual abuse and neglect, and other trauma history. Research evidence shows that patients who present with such risk factors are likely to experience future depressive episodes.217,80-86 Patients who have been treated for depression or suicidality in the past should continue to be monitored. PC clinicians should systematically evaluate adolescents at high risk for depression during health care visits (ie, well-child visits, urgent care visits). This systematic assessment should take place at least once a year, but frequent somatizers may need to be assessed more often.

Identification methods of youth at high risk may involve tools such as standardized written instruments, either generalized (eg, Guidelines for Adolescent Preventive Services and Strength and Difficulties Questionnaire) or specific emotional symptom checklists (eg, Beck Depression Inventory, Kutcher Adolescent Depression Scale). Although mnemonic-based interviews (eg, HEADSS: home, education/employment, activities, drugs, sexuality, suicide/depression) may also be used routinely during visits to guide the direct interview, systematic and scheduled use of psychometrically reliable and practical methods such as brief symptom checklists or validated depression scales are a preferred adjunct.

Assessment/Diagnosis

Recommendation 1: PC clinicians should evaluate for depression in adolescents at high risk as well as those who present with emotional problems as the chief complaint (grade of evidence: B; strength of recommendation: very strong). Clinicians should assess for depressive symptoms on the basis of diagnostic criteria established in the DSM-IV or International Classification of Diseases, 10th Revision (grade of evidence: B; strength of recommendation: very strong) and should use standardized depression tools to aid in the assessment (grade of evidence: A; strength of recommendation: very strong).

PC clinicians should probe for the presence of any of several depressive disorders, including MDD, dysthymia, and depression not otherwise specified by using systematic, rigorous assessment methods. Standardized instruments should be used to help with diagnosis but should not replace direct interview by the clinician.87-89 Because adolescents with depression may not be able to clearly identify depressed mood as their presenting complaint, providers need to be aware of common presenting symptoms that may signal MDD. These symptoms may include insomnia, weight loss, decline in academic functioning, family conflict, and other symptoms of depressive disorders.90 The Diagnostic and Statistical Manual for Primary Care can help PC clinicians distinguish between transient depressive responses and depressive disorders.

Recommendation 2: Assessment for depression should include direct interviews with the patients and families/caregivers (grade of evidence: B; strength of recommendation: very strong) and should include the assessment of functional impairment in different domains (grade of evidence: B; strength of recommendation: very strong) and other existing psychiatric conditions (grade of evidence: B; strength of recommendation: very strong).

Evidence of the core symptoms of depression and functional impairment should be obtained from the youth and from families/caregivers separately.92-94 The involvement of the family is critical in all phases of management and should be included in the assessment for depressive disorders. Family relationships also may affect the presentation of depression in adolescents. Cultural background of the patients and their families also must be considered during the assessment, because it can affect the presentation of core symptoms.95 Collateral information from other sources (such as teachers) may also be obtained to aid in the assessment. Given the high rates of comorbidities, clinicians should assess for the existence of comorbid conditions that may affect the diagnosis and treatment of the depressive disorder.2,17,96,97 These comorbidities may include 1 or more of the following conditions: substance abuse, anxiety disorder, attention-deficit/hyperactivity disorder, bipolar disorder, physical abuse, sexual abuse, and trauma. Instruments that assess for a range of common comorbid mental health conditions should be considered also. Clinicians should also assess for impairment in key areas of functioning including school, home, and peer settings.98 Subjective distress should be assessed also. Regardless of the diagnostic impression or any additional treatment plans, a safety assessment must be completed by the clinician (see recommendation 4 in “Initial Management of Depression”).
Initial Management of Depression

Recommendation 1: Clinicians should educate and counsel families and patients about depression and options for the management of the disorder (grade of evidence: C; strength of recommendation: very strong). Clinicians should also discuss limits of confidentiality with the adolescent and family (grade of evidence: D; strength of recommendation: very strong).

Management should be based on a plan developed with the understanding that depression is often a recurring condition. As seen in studies of depression interventions, families and patients need to be educated about the causes and symptoms of depression, impairments associated with it, and the expected outcomes of treatment.\textsuperscript{9,99-103} Information should be provided at a developmentally appropriate level, in a way that the patient and family can understand the nature of the condition and the management plan. Communication that is developmentally appropriate should facilitate the ability of parents and patients to work with the clinician to develop an effective and achievable treatment plan. To establish a strong therapeutic alliance, the clinician should also take into account cultural factors that may affect the diagnosis and management of this disorder.\textsuperscript{75} Clinicians should also be aware of the negative reactions of family members to a possible diagnosis of depression in the teen (ie, sadness, anger, denial). Sample materials are available in the GLAD-PC toolkit and include resources for patients and parents. Because the symptoms of depression can also affect many areas of an adolescent’s life, other ongoing partnerships may need to be established with personnel in schools and other settings (extracurricular activities). Confidentiality must also be discussed with the adolescent and his or her family. Adolescents and their families should be aware of the limits of confidentiality, including the need to involve parents or legal authorities when the risk of harm to the adolescent or others may be imminent. Clinicians should be aware of state laws regarding confidentiality (eg, see www.advocatesforyouth.org/publications/iag/confhlth.htm for additional information).

Recommendation 2: Clinicians should develop a treatment plan with patients and families (grade of evidence: C; strength of recommendation: very strong) and set specific treatment goals in key areas of functioning, including home, peer, and school settings (grade of evidence: D; strength of recommendation: very strong).

From studies of chronic disorders in youth, it is suggested that better adherence to treatment is associated with the identification and tracking of specific treatment goals and outcomes. Written action plans in asthma management have produced evidence for improved outcomes.\textsuperscript{102} If a patient presents with moderate-to-severe depression or has persistent depressive symptoms, treatment goals and outcomes should be identified and agreed upon, based on close collaboration with the patient and family at the time of treatment initiation. Treatment goals may include the establishment of a regular exercise routine, adequate nutrition, and regular meetings to resolve issues at home. In the adult depression literature, monitoring seems most effective when implemented through designated case managers who monitor patients’ clinical status and treatment-plan adherence.\textsuperscript{9} The benefits of such programs may be enhanced through the use of electronic medical charts and the development of patient registries.

Recommendation 3: The PC clinician should establish relevant links/collaboration with mental health resources in the community (grade of evidence: B; strength of recommendation: very strong), which may include patients and families who have dealt with adolescent depression and are willing to serve as resources to other affected adolescents and their family members (grade of evidence: D; strength of recommendation: very strong).

A major gap in the management of chronic disorders in young people is the lack of linkages between relevant services that make up the system of care for an individual youth.\textsuperscript{103} Furthermore, family-based interventions have been shown to help youth with mental illness.\textsuperscript{104} Therefore, establishing relevant links/collaboration with mental health resources in the local community, including peer support groups, advocacy groups, and traditional community- or hospital-based mental health services whenever these services are available, is essential to ensure timely and effective access to needed services.\textsuperscript{8,105} Such linkages may include prearranged agreement regarding referral, exchange of clinical information, points of contact, etc. Where appropriate (eg, rural areas), clinicians should also establish links with para-professionals who may provide the bulk of counseling and supportive services in underserved areas.

Recommendation 4: All management must include the establishment of a safety plan, which includes restricting lethal means, engaging a concerned third party, and developing an emergency communication mechanism should the patient deteriorate, become actively suicidal or dangerous to others, or experience an acute crisis associated with psychosocial stressors, especially during the period of initial treatment when safety concerns are highest (grade of evidence: C; strength of recommendation: very strong).

Suicidality, including ideation, behaviors, or attempts, is common among adolescents with depression. In studies of completed suicide, more than 50% of the victims had a diagnosis of depression.\textsuperscript{106} Therefore, clinicians who manage this disorder must develop an emergency communication mechanism for handling increased suicidality or acute crises. After assessing a suicidal patient for suicidality, the clinician must obtain information from a third party, assess that adequate adult supervision and support are available, have an adult agree to help remove lethal medications and firearms from the premises, warn the patient of the disin-
hbiting effects of drugs and alcohol, put contingency planning in place, and establish follow-up within a reasonable period of time.\textsuperscript{72,107} This plan should be developed with adolescents (and with their families/caregivers if possible) and should include a list of persons/services for the adolescent to contact in case of acute crisis or increased suicidality. The establishment of this plan is especially important during the period of diagnosis and initial treatment when safety concerns are highest. Clinicians may also work with schools to develop an emergency plan for all students who may experience an acute suicidal crisis. This global approach may prevent, in some instances, having to label a specific child suicidal when providers are merely trying to ensure that safety measures are in place in case the child decompensates. Components of a safety plan may also include a list of persons who are aware of the adolescents’ issues and will be able to assist if contacted during an acute crisis.

**DISCUSSION**

Although not definitive and subject to modification on the basis of ongoing accumulation of additional evidence, this part of the guidelines is intended to address the lack of recommendations regarding the screening, diagnosis, and initial management of depression in adolescents aged 10 to 21 years in PC settings in the United States and Canada. As such, these guidelines are intended to assist clinicians in family medicine, pediatrics, nursing, and internal medicine who may be the first (and sometimes only) clinicians to identify, manage, and possibly treat adolescent depression. These guidelines may also be helpful to allied health professionals who care for adolescents.

Although not all the steps involved in identifying, diagnosing, and initially managing the care for adolescent depression in PC have been (or even can be) subject to rigorous RCTs, there is sound reason to believe that existing tools and management protocols for adolescent depression can be applied in the PC setting. Although more research is needed, our review suggests that these components of the identification and initial management of adolescent depression in PC can be done. The recommendations were developed on the basis of areas that had at least “strong agreement” among experts. However, there were other controversial areas that were not addressed in these recommendations, such as universal screening. New emerging evidence may affect the inclusion of such areas in future iterations of these guidelines and the accompanying toolkit.

**Should These Guidelines Be Universally Deployed?**

One might question whether PC clinicians should identify and diagnose the problem of adolescent depression if the lack of psychiatric services prevents them from referring these youth.\textsuperscript{108} This caution notwithstanding, the increasingly prevailing recommendation is that as a minimum, PC clinicians should be provided the necessary guidance to support their initial management of adolescent depression.\textsuperscript{109,110} Nonetheless, because practitioners and their clinical practice settings vary widely in their degree of “readiness” in identifying and managing adolescent depression, it is likely that a good deal of time and flexibility will be required before these guidelines are adopted systematically or as a universal requirement. It is conceivable that integrated health care systems with electronic medical charts, tracking systems, and access to specialty mental health back-up and consultation will be most “ready” and able to fully implement the guidelines. The second part of the guidelines, the companion article,\textsuperscript{12} addresses the treatment of this disorder. Practices that do identify adolescent depression and have nowhere to refer the patients may benefit from the guidance offered in the next set of recommendations.

**Preparatory Steps**

Because the management of adolescent depression may constitute a new or major challenge for some PC practices, a number of important considerations should be kept in mind when preparing to implement the guidelines, given the findings from studies in the adult literature, input from our focus groups with clinicians, families and patients, and the experience of members of the GLAD-PC Steering Committee. Specifically, PC clinicians who manage adolescent depression should pursue (1) additional education regarding issues such as advances in screening, diagnosis, treatment, and follow-up, liability, consent, confidentiality, and billing, (2) practice and systems changes such as office staff training and “buy-in,” electronic medical charts, and automated tracking systems, whenever available, and (3) establishing linkages with mental health services.

Linkages with community mental health resources are necessary to both meet the learning needs of the PC clinician and facilitate consultation/referral of difficult cases. Practice and systems changes are useful in increasing clinicians’ capacity to ensure monitoring and follow-up of patients with depression. For example, staff training may help prioritize calls from adolescent patients who may not state the nature of their call. Specific tools and/or templates have been developed that offer examples of how to efficiently identify, monitor, track, and refer teens with depression. These materials are available in the GLAD-PC toolkit (available at www.glad-pc.org). The toolkit addresses how each of the recommendations might be accomplished without each practice necessarily having to “reinvent the wheel.”

**CONCLUSIONS**

Review of the evidence suggests that PC clinicians who have appropriate training and are attempting to deliver comprehensive health care should be able to identify and initiate management of adolescent depression. This
will likely require real changes in existing systems of care. As health care models such as the medical home indicate, comprehensive health care must include assessment and coordination of care for both physical and behavioral health. This first part of the guidelines for adolescent depression in PC may enable providers to pull together the current best evidence and attempt to initiate the best available high-quality care, even in instances in which they are not in a position to treat such youth. Mounting evidence suggests that pediatric providers can and should identify and coordinate depression care for their adolescent population.

APPENDIX: PART I TOOLKIT ITEMS

- Screening/assessment instruments (ie, Columbia Depression Scale)
- Information sheet on the developmental considerations in the diagnosis of depression
- Assessment algorithm/flow sheet (Fig 1)
- Fact sheet/family educational materials
- Educational materials on suicide prevention/safety planning

ACKNOWLEDGMENTS

We thank the Center for Substance Abuse Treatment (Substance Abuse and Mental Health Services Administration), the Josiah Macy, Jr Foundation, the New York State Office of Mental Health, the Lowenstein Foundation, the Center for the Advancement of Children’s Mental Health (Columbia University), Sunnybrook Health Sciences Centre (University of Toronto), the American Academy of Pediatrics (District II, New York chapters 1, 2 and 3), the New York Council on Child and Adolescent Psychiatry, the Children’s Health Forum (New York Academy of Medicine), the Kellogg Foundation, and the Civic Research Institute, Inc for financial support of the GLAD-PC project.

The GLAD-PC Steering Group consists of members of the GLAD-PC project team, steering committee, and official organizational liaisons. The GLAD-PC project team members are Peter S. Jensen, MD (project director, REACH Institute), Amy Cheung, MD (project coordinator, University of Toronto/Columbia University), Rachel A. Zuckerbrot, MD (project coordinator, Columbia University), Kareem Ghalib, MD (Columbia University), and Anthony Levitt, MD (project consultant, University of Toronto). The steering committee members are (listed alphabetically) Boris Birmaher, MD (Western Psychiatric Institute & Clinic, University of Pittsburgh), John Campo, MD (Ohio State University and Nationwide Children’s Hospital), Greg Clarke, PhD (Center for Health Research, Kaiser Permanente), Dave Davis, MD (University of Toronto), Angela Díaz, MD (Mount Sinai School of Medicine), Allen Dietrich, MD (Dartmouth Hitchcock Medical Center), Graham Emslie, MD (University of Texas Southwestern Medical School), Bernard Ewigman, MD (Department of Family Medicine, University of Chicago), Eric Fombonne, MD (McGill University), Sherry Glied, PhD (Columbia University), Kimberly Eaton Hoagwood, PhD (Office of Mental Health, New York State/Columbia University), Charles Homer, MD (National Initiative for Children’s Healthcare Quality), Danielle Laraque, MD (AAP New York Chapter 3, District II/Mount Sinai School of Medicine), Miriam Kaufman, MD (Hospital for Sick Children, University of Toronto), Kelly J. Kelleher, MD (Ohio State University), Stanley Kutzer, MD (Dalhousie Medical School), Michael Malus, MD (Department of Family Medicine, McGill University), James Perrin, MD (Massachusetts Medical School/Harvard Medical School), Harold Pincus, MD (Columbia University/New York State Psychiatric Institute), Brenda Reiss-Brennan, APRN (Intermountain Health), Diane Sacks, MD (Canadian Paediatric Society), Ruth E. K. Stein, MD (Forum for Child Health, New York Academy of Medicine, Albert Einstein College of Medicine), and Bruce Wastlick, MD, Baystate Health Systems, MA). The organizational liaisons are Angela Diaz, MD (American Academy of Pediatrics), Kelly Kelleher, MD (American Academy of Pediatrics), James Perrin, MD (American Academy of Pediatrics), Diane Sacks, MD (American Academy of Pediatrics/Canadian Paediatric Society), Bruce Wastlick, MD (American Medical Association), David Fassler, MD (American Academy of Child and Adolescent Psychiatry), Eric Fombonne, MD (Canadian Academy of Child Psychiatry and Canadian Psychiatric Association), James McIntyre, MD (American Psychiatric Association), Judy Garber, PhD (American Psychological Association), Vicky Wolfe, PhD (Canadian Psychological Association), Michael Malus, MD (College of Family Medicine of Canada), Johanne Renaud, MD (Canadian Association for Adolescent Health), Debbie Ebner, PhD (Society for Adolescent Medicine), Stanford Friedman, MD (Society for Developmental and Behavioral Pediatrics), Terry Stancin, PhD (Society for Developmental and Behavioral Pediatrics), Kathryn Salisbury, PhD (Mental Health Association of New York City), Michael Faenza, MSSW (National Mental Health Association), Susan Bergeson (Depression and Bipolar Support Alliance), Darcy Gruttadaro (National Alliance on Mental Illness), Sandra Spencer (Federation of Families for Children’s Mental Health), and Elizabeth Hawkins-Walsh, DNSc, CPNP (National Association for Pediatric Nurse Practitioners).

In addition to members of the GLAD-PC Steering Group, we also acknowledge the GLAD-PC conference attendees: Perry Adler, PhD; Joan Asarnow, PhD; Loretta Young Au, MD, FAAP; Abraham Bartell, MD; Tamar Bauer; Rachel Bergeson, MD; Blanche Benenson, MD; Linda Theil Cahill, MD; Wayne Cannon, MD; Marie Barone Casalino, MD; Sonia Chehil, MD; Joseph Cra-
Preparation for Managing Depression in Primary Care
Preparation through increased training, establishing mental health linkages, and increasing the capacity of practices to monitor and follow-up with patients with depression.

Youth presents to clinic for urgent care or health maintenance visit

Early Identification
Systematically identify high-risk youth

(1) Stop assessment
(2) Repeat surveillance as needed

If low risk
If high risk or presenting with emotional issues as chief complaint

Youth or family presents with emotional issues as chief complaint

Assessment
(1) Assess with systematic depression assessment tool
(2) Interview patient and parent to assess for depression and other psychiatric disorders with DSM-IV or ICD-10 criteria
(3) Assess for safety/suicide risk

Evaluation Negative for Depression, but positive for other MH conditions

(1) Refer to other treatment guidelines;
(2) Evaluate for depression at future visits
(3) Book for follow-up visit.

Evaluation Positive for Depression: MILD, MODERATE, SEVERE, or Depression with COMORBIDITIES

(1) Evaluate safety and establish safety plan.
(2) Evaluate severity of depression symptoms. a
(3) Patient/Family Education. b
(4) Develop treatment plan based on severity-review diagnosis and treatment options with patient/family.

If psychotic or suicidal

Refer to Crisis or Emergency Services (may include subsequent referral to inpatient treatment)

FIGURE 1
Clinical assessment flowchart.

a See Guidelines Part I for definition of mild, moderate, and severe depression. Please consult toolkit for methods available to aid clinicians to distinguish between mild, moderate, and severe depression.

b Psychoeducation, supportive counseling, facilitate parental and patient self-management, refer for peer support and regular monitoring of depressive symptoms and suicidality.
mer, MD; Cathryn Cunningham, MD; Wendy Davis, MD; Carolyn Dewa, PhD; Benard P. Dreyer, MD; M. Flament, MD; D. Clare Fried, MD; William Gardner, PhD; Neville Golden, MD; Catherine Goodfellow, MD; Myla Harrison, MD; Sarah Horwitz, PhD; Barbara Hull; Daniel Hyman, MD; Maria Kovacs, PhD; Deborah Launer; Susan Lippert Levitzky, MD; Robert Lubarsky, MD; Christopher P. Lucas, MD; Wanda McCoy, MD; Thomas McNerny, MD; Jessica Mass-Levitt, PhD; Margaret McHugh, MD; Laura Muñoz, PhD; Gwen Nilsson, MD; Elizabeth Pappadopulos, PhD; Matthew Perkins, MD; Ellen Perrin, MD; Kathryn Salisbury; Marcie Beth Schneider, MD; Warren Seigel, MD; Tamara Singer, MD; Karen Soren, MD; L. Read Sulik, MD; Kristin Trautman; John Van Gorder; Benedetto Vitiello, MD; Robin Weersing, MD; Myrna Weissmann, PhD; Eric Weiselberg, MD; Karen Soren, MD; L. Read Sulik, MD; Kristin Trautman; John Van Gorder; Benedetto Vitiello, MD; Robin Weersing, MD; Myrna Weissmann, PhD; Eric Weiselberg, MD; Mark L. Wolraich, MD; and Alan Wong, MD.

REFERENCES

32. Smith MS, Mitchell J, McCauley EA, Calderon R. Screening
41. Reynolds WM. Reynolds Adolescent Depression Scale. Lutz, FL: Psychological Assessment Resources; 1987
54. La Greca AM, Bearman KJ. Commentary: if “an apple a day keeps the doctor away,” why is adherence so darn hard? J Pediatr Psychol. 2001;26:279–282
56. Stroul BA, Friedman R, eds. A System of Care for Children and Youth With Severe Emotional Disturbances. Washington, DC: Georgetown University Child Development Center, CASSP, Technical Assistance Center; 1986
74. Shah S, Hoffman RE, Wake L, Marine WM. Adolescent sui-


103. Stroul BA, Friedman RM. A System of Care for Children and Youth With Severe Emotional Disturbances. Washington, DC: Georgetown University Child Development Center, CASSP Technical Assistance Center


Guidelines for Adolescent Depression in Primary Care (GLAD-PC): II. Treatment and Ongoing Management

Amy H. Cheung, MD, Rachel A. Zuckerbrot, MD, Peter S. Jensen, MD, Kareem Ghalib, MD, Danielle Laraque, MD, Ruth E. K. Stein, MD, and the GLAD-PC Steering Group

OBJECTIVES. To develop clinical practice guidelines to assist primary care clinicians in the management of adolescent depression. This second part of the guidelines addresses treatment and ongoing management of adolescent depression in the primary care setting.

METHODS. Using a combination of evidence- and consensus-based methodologies, guidelines were developed in 5 phases as informed by (1) current scientific evidence (published and unpublished), (2) a series of focus groups, (3) a formal survey, (4) an expert consensus workshop, and (5) revision and iteration among members of the steering committee.

RESULTS. These guidelines are targeted for youth aged 10 to 21 years and offer recommendations for the management of adolescent depression in primary care, including (1) active monitoring of mildly depressed youth, (2) details for the specific application of evidence-based medication and psychotherapeutic approaches in cases of moderate-to-severe depression, (3) careful monitoring of adverse effects, (4) consultation and coordination of care with mental health specialists, (5) ongoing tracking of outcomes, and (6) specific steps to be taken in instances of partial or no improvement after an initial treatment has begun. The strength of each recommendation and its evidence base are summarized.

CONCLUSIONS. These guidelines cannot replace clinical judgment, and they should not be the sole source of guidance for adolescent depression management. Nonetheless, the guidelines may assist primary care clinicians in the management of depressed adolescents in an era of great clinical need and a shortage of mental health specialists. Additional research concerning the management of youth with depression in primary care is needed, including the usability, feasibility, and sustainability of guidelines and determination of the extent to which the guidelines actually improve outcomes of youth with depression.

**Key Words**
adolescents, depression, primary care, guidelines

**Abbreviations**
PC—primary care
MDD—major depressive disorder
AACAP—American Academy of Child and Adolescent Psychiatry
PCP—primary care provider
GLAD-PC—Guidelines for Adolescent Depression in Primary Care
FDA—Food and Drug Administration
RCT—randomized, controlled trial
CES-DC—Center for Epidemiological Studies Depression Scale for Children
MAOI—monoamine oxidase inhibitor
CBT—cognitive behavioral therapy
SSRI—selective serotonin reuptake inhibitor
IPT—interpersonal psychotherapy
CI—confidence interval
STUDIES HAVE SHOWN that up to 9% of teenagers meet criteria for depression at any one time, with as many as 1 in 5 teens having a history of depression at some point during adolescence.1–5 In primary care (PC) settings, point prevalence rates are likely higher, with rates up to 28%.6–10 Taken together, epidemiologic and PC-specific studies have suggested that despite relatively high rates, major depressive disorder (MDD) in youth is underidentified and undertreated in PC settings.11

Because of barriers to adolescents receiving specialty mental health services, only a small percentage of depressed adolescents are treated by mental health professionals.12 As a result, PC settings have become the de facto mental health clinics for this population, although most PC clinicians feel inadequately trained, supported, or reimbursed for the management of this disorder.13–18 Although MDD management guidelines have been developed for specialty care settings (eg, see the American Academy of Child and Adolescent Psychiatry [AACP] practice parameters19) or for related problems such as suicidal ideation or attempts,20 it is clear that significant practice and clinician differences exist between the primary and specialty care settings that do not allow a simple transfer of guidelines from one setting to another.

Recognizing this gap in clinical guidance for PC providers (PCPs), a group of researchers from the United States and Canada established the Guidelines for Adolescent Depression in Primary Care (GLAD-PC), a North American collaborative, to develop guidelines for the management of adolescent depression in the PC setting. The development process of GLAD-PC is described in detail in our companion report.21 This article describes the recommendations regarding treatment, ongoing management, and follow-up along with the supporting empirical evidence for those recommendations. Our companion article provides the corresponding evidence and resulting recommendations for depression identification, assessment and diagnosis, and initial management before the formal onset of specific treatments.21

METHODS
A full description of the methodology used for the development of GLAD-PC is included in our companion article.21 In brief, the expert collaborative used a mix of qualitative (focus groups, expert consensus) and quantitative (survey, literature reviews) methods to inform the development of GLAD-PC. In view of space limitations, only the results of the literature reviews regarding available evidence pertaining to treatment, ongoing management, and follow-up procedures are presented in this article.

Literature Reviews
The literature-review approach used for all of the reviews was as follows. First, the GLAD-PC team identified the existence of high-quality, previously published, systematic evidence-based reviews that met the following criteria: (1) explicit definition of search terms and years covered; (2) exhaustive search of Medline; (3) reading of abstracts to determine relevance, followed by review of entire articles from relevant abstracts; (4) restriction to English-language journals only; (5) restriction to empirical articles; and (6) identification of any otherwise omitted citations from the reference sections from key reviews. In areas where there were no carefully executed and well-described systematic literature reviews had been recently conducted (ie, Food and Drug Administration [FDA], Cochrane), the GLAD-PC team conducted a systematic review for primary studies for each area by using Medline (from inception to 2004/2005) based on the 5 criteria described above.

Three literature reviews were conducted for the GLAD-PC recommendations presented in this article: (1) nonspecific psychosocial interventions in pediatric PC,22 (2) antidepressant treatment,23 and (3) the use of psychotherapy. For the first review, Stein et al22 searched the literature (Medline, PsycINFO, and the Cochrane database) for articles that examined evidence for psychosocial interventions delivered in the PC setting. The reference lists of all relevant articles were searched for additional studies. In addition, experts in the field were consulted to identify additional studies. Given the paucity of randomized, controlled trials (RCTs) identified earlier in a review by Bower et al24 in 2001, studies with simple before-and-after comparisons were also included.

In the second review, we examined the efficacy and safety of antidepressant medications in the pediatric population (aged 7–18). The studies were identified in 2 stages. Given the thorough reanalyses of safety data on both published and unpublished clinical trials completed by the FDA, all RCTs included in the FDA safety report were reviewed.25 Second, to ensure that additional studies not reported to the FDA were not missed, Medline and PsycINFO were searched. For a full description of the review, please refer to the published review.23

In the final review, we searched the literature for depression trials that examined the efficacy of psychotherapy. The search included all forms of psychotherapy including both individual and group-based therapies. We not only identified individual studies but also high-quality systematic reviews given the extensive empirical literature in this area. Additional details on each of these searches, including search terms, number of abstracts selected, etc, are available from the authors on request.

Expert Consensus
Expert consensus was reached through 2 stages. First, expert participants completed a survey regarding adolescent depression management. Subsequently, the expert participants then met in a 2-day workshop to review the survey results to reach consensus on key issues regard-
ing identification and treatment of adolescent depression in PC. Overall, the guidelines only included recommendations that the experts agreed are highly appropriate and “first-line” practices.

RESULTS: LITERATURE REVIEWS

Psychosocial Interventions in PC

On behalf of the GLAD-PC team, Stein et al.22 reviewed the evidence for the efficacy of PC-delivered psychosocial interventions. The studies identified were divided into 2 categories of evidence: direct and indirect. Direct evidence included data from studies that evaluated interventions specific for adolescent depression, and indirect evidence included data from studies that examined PC interventions for adults with depression and PC interventions for other psychosocial problems in the pediatric population. Additional details about the review can be found in the Stein et al report.22

The literature review identified 4 articles that focused on depression interventions in the PC setting for adolescents,26–29 all of which showed positive outcomes for the PC-delivered interventions. Walker et al conducted an RCT to evaluate the effectiveness of a PC-delivered consultation intervention that involved teens from 8 general practices in Britain. Teens were invited to discuss their health concerns with PC nurses who provided individual consultations.22,26 The PC nurses also offered mental health referrals when appropriate. In those with high Center for Epidemiological Studies Depression Scale for Children (CES-DC) scores, those who were randomly assigned to PC nurse consultation had lower CES-DC scores on follow-up than adolescents with high CES-DC scores who were not randomly assigned to the consultation. The results suggest that PC-delivered intervention may be helpful in reducing depressive symptomology as measured by the CES-DC.

Stein et al.22 also identified 6 additional studies that focused on PC counseling. 4 that focused on improved patient outcomes for adult depression, and 2 that focused on improved parent-child relationships in postpartum depression. The adult depression in PC literature has shown that psychosocial support by a physician, nurse, or other staff, in the context of 15-minute problem-solving therapy, improves outcomes in depressed adults.30–32

With the pediatric PC literature, findings revealed that authors of previous studies have attempted to train pediatricians in various types of counseling such as anticipatory guidance or preventive counseling in a number of disorder areas. Before the Stein et al review,22 the most recently published systematic review on the impact of psychosocial interventions in pediatric PC was conducted by Bower et al.24 These investigators reviewed 25 studies in pediatric PC that demonstrated tremendous variability in the problems treated, clinician interven-

tions, and outcomes studied.24 As noted by Stein et al,22 most studies did not compare the intervention group to a control group, and those that were RCTs did not provide enough information to judge the design. Thus, one must be cautious in stating that the “innovative” PC interventions were superior to usual care, although some positive effects were found. However, as shown in a review by Bass et al,13 18 studies have demonstrated positive effects of injury-prevention counseling in pediatric PC, and Stein et al22 reviewed additional studies that suggested that modest educational counseling performed by pediatric PC staff can be useful.

Antidepressant Treatment

The treatment review for antidepressant safety and efficacy included RCTs of antidepressants in youth under the age of 19 with depression. This review has also been published elsewhere.23 This GLAD-PC–initiated review identified 8 peer-reviewed articles in this area, including 4 trials with fluoxetine,34–37 1 with sertraline,38 1 with citalopram,38 1 with paroxetine,39 and 1 with venlafaxine.40 Older antidepressants (ie, monoamine oxidase inhibitors [MAOIs], tricyclic antidepressants) were not included in our review because the current controversy and the recent FDA review only involved newer classes of antidepressants and because of the known lack of efficacy (ie, tricyclic antidepressants) and clinical trials data (ie, MAOIs) for other classes of older antidepressants.30 Finally, because of the continuing controversy around the disclosure of unpublished clinical research data, unpublished studies included in the FDA review were also reviewed. There were no completed yet unpublished studies identified that were excluded in the FDA analyses. For additional details regarding this review, see the Cheung et al report.25

Overall, both individual clinical trial evidence and evidence from systematic reviews support the use of antidepressants in adolescents with MDD. Bridge et al41 conducted a meta-analysis of the clinical trials data and calculated the numbers needed to treated and numbers needed to harm. They concluded that 6 times more teens would benefit from treatment with antidepressants than would be harmed.41 In reviewing the individual studies, the percentage of subjects who responded to antidepressants ranged from 47% to 69% and 33% to 57% for those on placebo (see Table 1). The majority of these studies found a significant difference between those on medication versus those on placebo. Overall, fluoxetine has had the largest number of studies with positive results, whereas paroxetine has had the largest number of studies with negative results.34–37,39,42,43 However, methodologic differences may have played a role in these differing results.23 The largest study, the Treatment for Adolescent Depression Study, involved subjects who were randomly assigned to receive placebo, cognitive behavioral therapy (CBT) alone, fluoxetine alone, or
CBT with fluoxetine. Subjects assigned to receive CBT with fluoxetine or fluoxetine alone showed significantly greater improvement in their depressive symptoms compared with those who received placebo or were treated with CBT alone (also see “Cognitive Behavioral Therapy”).

Finally, available evidence from several large RCTs suggests that adverse effects do emerge in depressed youth who are treated with antidepressants. Adverse effects (ie, nausea, headaches, behavioral activation, etc) occur in up to 93% of the subjects treated with these medications and in up to 75% of those treated with placebo when subjects are asked about specific adverse effects. Therefore, routine monitoring of the development of adverse events is critical for depressed youth who are treated with antidepressants.

Authors of several recent studies have used population data to evaluate the risks versus benefits of prescribing antidepressants. Olfson et al focused specifically on youth aged 10 to 19 years; their study revealed decreased suicide rates in geographic areas where the rates of newer antidepressant prescriptions are increasing. Gibbons et al reported similar findings when they conducted a study of children aged 5 to 14. Several other studies have focused on general populations that included significant numbers of children and adolescents. Although one Australian study did identify a link between increased prescription rates of newer antidepressants and increased suicide rates in adolescents and young adults aged 15 to 24, other American and international studies have indicated an inverse relationship between rates of selective serotonin reuptake inhibitor (SSRI) prescriptions and rates of suicide in adolescent populations.

Still other studies have used large databases to carry out naturalistic studies of possible associations between antidepressant use and suicidality. In the only 1 of these studies that focused exclusively on youth, Valuck et al conducted a propensity-adjusted retrospective co-

<table>
<thead>
<tr>
<th>Medication</th>
<th>Drug, %</th>
<th>Placebo, %</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluoxetine (March et al [2004])a</td>
<td>56</td>
<td>33</td>
<td>.02</td>
</tr>
<tr>
<td>Fluoxetine (Emslie et al [1997])</td>
<td>52</td>
<td>36.8</td>
<td>.03</td>
</tr>
<tr>
<td>Fluoxetine (Emslie et al [2002])</td>
<td>61</td>
<td>35</td>
<td>.001</td>
</tr>
<tr>
<td>Paroxetine (Keller et al [2001])b</td>
<td>66</td>
<td>48</td>
<td>.02</td>
</tr>
<tr>
<td>Paroxetinec</td>
<td>69</td>
<td>57.3</td>
<td>NS</td>
</tr>
<tr>
<td>Paroxetinec</td>
<td>65</td>
<td>46</td>
<td>.005</td>
</tr>
<tr>
<td>Citalopram (Wagner et al [2004])</td>
<td>47</td>
<td>45</td>
<td>NS</td>
</tr>
<tr>
<td>Sertraline (Wagner et al [2003])</td>
<td>63</td>
<td>53</td>
<td>.05</td>
</tr>
<tr>
<td>Escitalopram (Wagner et al [2007])</td>
<td>63</td>
<td>52</td>
<td>NS</td>
</tr>
</tbody>
</table>

NS indicates not significant.

a Fluoxetine alone compared with placebo.
b Paroxetine compared with placebo.
c GlaxoSmithKline, unpublished data.


TABLE 1 Response Rates in RCTs of Antidepressants Based on Clinical Global Impression

<table>
<thead>
<tr>
<th>Medication</th>
<th>Drug, %</th>
<th>Placebo, %</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluoxetine (March et al [2004])a</td>
<td>56</td>
<td>33</td>
<td>.02</td>
</tr>
<tr>
<td>Fluoxetine (Emslie et al [1997])</td>
<td>52</td>
<td>36.8</td>
<td>.03</td>
</tr>
<tr>
<td>Fluoxetine (Emslie et al [2002])</td>
<td>61</td>
<td>35</td>
<td>.001</td>
</tr>
<tr>
<td>Paroxetine (Keller et al [2001])b</td>
<td>66</td>
<td>48</td>
<td>.02</td>
</tr>
<tr>
<td>Paroxetinec</td>
<td>69</td>
<td>57.3</td>
<td>NS</td>
</tr>
<tr>
<td>Paroxetinec</td>
<td>65</td>
<td>46</td>
<td>.005</td>
</tr>
<tr>
<td>Citalopram (Wagner et al [2004])</td>
<td>47</td>
<td>45</td>
<td>NS</td>
</tr>
<tr>
<td>Sertraline (Wagner et al [2003])</td>
<td>63</td>
<td>53</td>
<td>.05</td>
</tr>
<tr>
<td>Escitalopram (Wagner et al [2007])</td>
<td>63</td>
<td>52</td>
<td>NS</td>
</tr>
</tbody>
</table>

NS indicates not significant.
a Fluoxetine alone compared with placebo.
b Paroxetine compared with placebo.
c GlaxoSmithKline, unpublished data.


TABLE 1 Response Rates in RCTs of Antidepressants Based on Clinical Global Impression

<table>
<thead>
<tr>
<th>Medication</th>
<th>Drug, %</th>
<th>Placebo, %</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluoxetine (March et al [2004])a</td>
<td>56</td>
<td>33</td>
<td>.02</td>
</tr>
<tr>
<td>Fluoxetine (Emslie et al [1997])</td>
<td>52</td>
<td>36.8</td>
<td>.03</td>
</tr>
<tr>
<td>Fluoxetine (Emslie et al [2002])</td>
<td>61</td>
<td>35</td>
<td>.001</td>
</tr>
<tr>
<td>Paroxetine (Keller et al [2001])b</td>
<td>66</td>
<td>48</td>
<td>.02</td>
</tr>
<tr>
<td>Paroxetinec</td>
<td>69</td>
<td>57.3</td>
<td>NS</td>
</tr>
<tr>
<td>Paroxetinec</td>
<td>65</td>
<td>46</td>
<td>.005</td>
</tr>
<tr>
<td>Citalopram (Wagner et al [2004])</td>
<td>47</td>
<td>45</td>
<td>NS</td>
</tr>
<tr>
<td>Sertraline (Wagner et al [2003])</td>
<td>63</td>
<td>53</td>
<td>.05</td>
</tr>
<tr>
<td>Escitalopram (Wagner et al [2007])</td>
<td>63</td>
<td>52</td>
<td>NS</td>
</tr>
</tbody>
</table>

NS indicates not significant.
a Fluoxetine alone compared with placebo.
b Paroxetine compared with placebo.
c GlaxoSmithKline, unpublished data.


TABLE 1 Response Rates in RCTs of Antidepressants Based on Clinical Global Impression

<table>
<thead>
<tr>
<th>Medication</th>
<th>Drug, %</th>
<th>Placebo, %</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluoxetine (March et al [2004])a</td>
<td>56</td>
<td>33</td>
<td>.02</td>
</tr>
<tr>
<td>Fluoxetine (Emslie et al [1997])</td>
<td>52</td>
<td>36.8</td>
<td>.03</td>
</tr>
<tr>
<td>Fluoxetine (Emslie et al [2002])</td>
<td>61</td>
<td>35</td>
<td>.001</td>
</tr>
<tr>
<td>Paroxetine (Keller et al [2001])b</td>
<td>66</td>
<td>48</td>
<td>.02</td>
</tr>
<tr>
<td>Paroxetinec</td>
<td>69</td>
<td>57.3</td>
<td>NS</td>
</tr>
<tr>
<td>Paroxetinec</td>
<td>65</td>
<td>46</td>
<td>.005</td>
</tr>
<tr>
<td>Citalopram (Wagner et al [2004])</td>
<td>47</td>
<td>45</td>
<td>NS</td>
</tr>
<tr>
<td>Sertraline (Wagner et al [2003])</td>
<td>63</td>
<td>53</td>
<td>.05</td>
</tr>
<tr>
<td>Escitalopram (Wagner et al [2007])</td>
<td>63</td>
<td>52</td>
<td>NS</td>
</tr>
</tbody>
</table>

NS indicates not significant.
a Fluoxetine alone compared with placebo.
b Paroxetine compared with placebo.
c GlaxoSmithKline, unpublished data.

Psychotherapy

The final review conducted examined the efficacy of psychotherapy such as CBT, interpersonal psychotherapy (IPT), and nonspecific interventions such as counseling and support. Through our search, we were able to identify both individual studies and several high-quality meta-analyses/reviews that were recently conducted to examine the efficacy of psychotherapy in adolescent depression. A full description of the review is available from the authors on request.

Cognitive Behavioral Therapy

In 1998, Reinecke et al and Harrington et al conducted reviews of CBT trials and found improved outcomes. In addition to the above-mentioned meta-analytical studies, several systematic narrative reviews of CBT studies have been conducted. The most recent and comprehensive of these reviews was conducted by Compton et al and included 12 studies published between 1990 and 2002. Although some of these studies showed negative results, Compton et al conclude that, in sum, they provided solid evidence of the effectiveness of CBT conducted by trained therapists for mild-to-moderate depression.

The effectiveness of CBT for adolescents with moderately severe depression was evaluated recently in the multicenter Treatment for Adolescents With Depression Study, which randomly assigned 439 depressed 12- to 17-year-olds to treatment with CBT, fluoxetine, CBT plus fluoxetine, or placebo. According to Clinical Global Impressions severity scores, the post-treatment response rate to 15 sessions of CBT over 12 weeks (43.2% [95% confidence interval [CI]: 34–52]) was not significantly different (P = .40) from placebo (34.8% [95% CI: 26–44]). The authors attributed this relatively low response rate, in part, to the fact that the study population suffered from more severe and chronic depression than participants in previous studies and to a high rate of psychiatric comorbidity in their study participants. Along with the fairly robust placebo-response rate, it is also possible that the nonspecific therapeutic aspects of this medication management could have successfully competed with the specific effects of the CBT intervention. As a consequence, one cannot and should not conclude that CBT was ineffective.

Although the Compton et al review of studies pro-
vided evidence for the efficacy of CBT in specialty mental health clinics for adolescents with mild-to-moderate depression, more recent studies have helped determine the effectiveness of CBT in “real-world” situations. In 2003, Puskar et al57 evaluated whether a group CBT intervention conducted in a high school by a masters-level nurse could improve depressive symptoms among 89 rural students with Reynold’s Adolescent Depression Screen scores of ≥60. The 46 students who completed 10 weekly CBT group sessions had significantly better mean depressive scores immediately after treatment and at 6 months than those (n = 43) who were randomly assigned to receive usual care. In contrast, Kerfoot et al58 studied the impact of training social workers in CBT methods (versus treatment as usual) on 52 depressed youth. The study showed no differences in depression scores across the 2 interventions, which was partially attributed to high drop-out rates.58

In yet another study with more difficult adolescents, Rohde et al59 recently assessed the effectiveness of CBT in treating adolescents with comorbid MDD and conduct disorder by recruiting 13- to 17-year-olds (N = 93). After randomly assigning them to a CBT-based “Coping With Depression” course or a life skills tutoring control condition, 39% of the adolescents who “completed” the CBT course recovered compared with only 19% of the adolescents who participated in the life skills tutoring control group (odds ratio: 2.66; 95% CI: 1.03–6.85).

Finally, in the Youth Partners-in-Care study, Asarnow et al27 evaluated the effectiveness of a quality improvement intervention that involved increasing access of PC clinicians and depressed youth to CBT and antidepressant medication. Participants (N = 418) were randomly assigned to usual care “enhanced” by an education intervention or the quality improvement intervention.27 At the study’s 6-month end point, subjects in the intervention group were significantly improved according to the study’s 2 primary outcome variables as measured by the Center for Epidemiological Study Depression Scale. The intervention group had lower Center for Epidemiological Study Depression Scale scores and fewer youth scored in the severe range at the end of the study. Given the fact that the intervention and usual-care groups differed significantly only in their use of CBT (53% vs 36%, respectively; OR: 2.2; 95% CI: 1.3–3.9; P = .007), much of the intervention groups’ improvement can be attributed to the availability of this treatment modality for patients screened and identified in PC settings.

Interpersonal Therapy
In terms of IPT, only a handful of studies have been conducted. First, Mulson et al38 assigned 48 depressed adolescents to IPT for adolescents (IPT-A) or clinical monitoring. Those who received IPT-A reported fewer depressive symptoms and improved overall functioning. In the Rossello and Bernal60 study, 71 depressed Puerto Rican adolescents were randomly assigned to receive IPT, CBT, or be placed on a waiting list. After 12 weeks, both IPT- and CBT-treated adolescents reported significantly fewer depressive symptoms. In the most recent study, 63 depressed adolescents (any depressive disorder) were randomly assigned to receive either 16 weeks of IPT-A or a treatment-as-usual condition (supportive counseling).61 Subjects who were treated with IPT-A showed significantly greater symptom reduction and improved overall functioning.

GUIDELINES
Each of the recommendations listed below was graded on the basis of the level of supporting research evidence from the literature and the extent to which experts agreed that it is highly appropriate in PC. The level of supporting evidence for each recommendation is based on the Oxford Centre for Evidence-Based Medicine grades of evidence (A–D) system (see www.cebm.net/levels_of_evidence.asp).

Recommendation strength based on expert consensus was rated in 4 categories: very strong (>90% agreement), strong (>70% agreement), fair (>50% agreement), and weak (<50% agreement). The recommendations in the guidelines were developed only in areas of management that had at least “strong agreement” among experts (see Fig 1 for the treatment algorithm).

Treatment
Recommendation 1: After initial diagnosis, in cases of mild depression, clinicians should consider a period of active support and monitoring before starting other evidence-based treatment (grade of evidence: B; strength of recommendation: very strong).

After a preliminary diagnostic assessment, in cases of mild depression, clinicians should consider a period of active support and monitoring before recommending treatment (from 6 to 8 weeks of weekly or biweekly visits for active monitoring). Evidence from RCTs of antidepressants and CBT show that a sizable percentage of patients respond to nondirective supportive therapy and regular symptom monitoring.34-40,62-65 However, if symptoms persist, treatment with antidepressants or psychotherapy should be offered. Active support and monitoring is also essential for cases in which depressed patients and/or their families/caregivers refuse other treatments. Active support and counseling for adolescents by pediatric PC clinicians have been evaluated for several different disorders including substance abuse and sleep disorders.22

Furthermore, expert opinion based on extensive clinical experience and qualitative research with families, patients, and clinicians indicate that these strategies are a crucial component of management by PC clinicians. For additional guidance on how to provide active sup-
For moderate or severe cases, the clinician should recommend treatment, crisis intervention (as indicated), and mental health consultation immediately without a period of active monitoring.

Recommendation 2: If a PC clinician identifies an adolescent with moderate or severe depression or complicating factors/conditions such as coexisting substance abuse or psychosis, consultation with a mental health specialist should be considered (grade of evidence: C; strength of recommendation: strong). Appropriate roles and responsibilities for ongoing management by the PC and mental health clinicians should be communicated and agreed upon (grade of evidence: C; strength of rec-
ommendation: strong). The patient and family should be consulted and approve the roles of the PC and mental health professionals (grade of evidence: D; strength of recommendation: strong).

In adolescents with severe depression or comorbidities such as substance abuse, clinicians should consider consultation with mental health professionals and refer to such professionals when deemed necessary. In cases of moderate depression with or without comorbid anxiety, clinicians should consider consultation by mental health and/or treatment in the PC setting. Although the access barriers to mental health services need to be addressed by policy makers to make necessary mental health consultations more feasible, available, and affordable in underserviced areas, clinical judgment must prevail in the meantime; thus, the need for consultation should be based on the clinician’s judgment. PC clinicians must also take into consideration the treatment preferences of patients/families, the severity and urgency of the case presentation, and the physician’s level of training and experience.

Active support and treatment should also be started in cases in which there is a lengthy waiting list for mental health services. Once a referral is made, the PC clinician must remain involved in the follow-up. In particular, roles and responsibilities should be agreed upon between the PC clinician and mental health clinician(s), including the designation of case coordination responsibilities.66–72

Recommendation 3: PC clinicians should recommend scientifically tested and proven treatments (ie, psychotherapies such as CBT or IPT and/or antidepressant treatment such as SSRIs) whenever possible and appropriate to achieve the goals of the treatment plan (grade of evidence: A; strength of recommendation: very strong).

After providing education and support to the patient and family, the range of effective treatment including medications, psychotherapies, and family support options should be considered. The patient and family should be assisted to arrive at a treatment plan that is both acceptable and implementable while taking into account their preferences and availability of treatment services. The treatment plan should be customized according to the severity of disease, risk of suicide, and the existence of comorbid conditions. The GLAD-PC toolkit will provide more detailed guidance around the factors that may influence a treatment choice (ie, a patient with psychomotor retardation may not be able to actively engage in psychotherapy). The management of depression in youth is an emerging field, and new treatments may become available. However, common-sense approaches such as the prescription of physical exercise and adequate nutrition should also be used in the management of these patients.

As an aside, the majority of CBT and IPT studies that included patients with MDD also included patients with depression not otherwise specified, subthreshold depressive symptoms, or dysthymic disorder. In contrast, medication RCTs for depression in adolescents generally only included subjects with MDD. Thus, although these guidelines address the treatment of depression generally, medication-specific guidelines apply only to fully expressed MDD.

Psychotherapies
Both CBT and IPT have been adapted to address depression in adolescents and have been shown to be effective in treating adolescents with MDD in tertiary care and in community settings.28,61,73 CBT has been used in the PC setting with positive preliminary results.27,29 However, the results of a recent RCT demonstrated superior efficacy of combination therapy (medication and CBT) versus CBT alone.36 For a brief description of the 2 therapies, see Table 2.

Antidepressant Treatment
Previous research has shown that up to 25% of pediatric PC clinicians and 42% of family physicians in the United States had recently prescribed SSRIs for more than 1 adolescent under the age of 18.13 When indicated by clinical presentation (clear diagnosis of MDD with no comorbid conditions) and patient/family preference, an SSRI should be used. The selection of the specific SSRI should be based on the optimum combination of safety and efficacy data. The patient and family should be informed about the possible adverse effects (clinicians may use a checklist) including possible switch to mania or the development of behavioral activation or suicidal behavior. Once the antidepressant is started, and if tolerated, the clinician should ensure an adequate trial up to the maximum dose and duration.

Table 3 lists recommended antidepressants and dosages. Required fields are marked with an asterisk (*).
ages for use in youth with depression. These recommendations are based on the expert survey results and were also reviewed by our expert consensus panel. Generally, the effective dosages for antidepressants in adolescents are lower than would be found in adult guidelines. Note that only fluoxetine has been approved by the FDA for use in children and adolescents with depression. Clinicians should know the potential drug interactions with SSRIs. Further information on the use of antidepressants are described in the GLAD-PC toolkit. In addition, all SSRIs, with the exception of fluoxetine, should be slowly tapered when discontinued because of the risk of withdrawal effects. Details regarding the initial selection of a specific SSRI and possible reasons for initial drug choice can be found in the GLAD-PC toolkit.

Contact (either in person or by telephone with either the clinician or member of the clinical staff) should take place after the initiation of treatment to review the patient’s and family’s understanding of and adherence to the treatment plan. Issues such as the current status of the patient and the patient’s/family’s access to educational materials regarding depression should be discussed during follow-up conversations. For relevant educational resources for patients and/or families, refer to the GLAD-PC toolkit.

**Recommendation 4: PC clinicians should monitor for the emergence of adverse events during antidepressant treatment (SSRIs)** (grade of evidence: B; strength of recommendation: very strong).

Recent reanalyses of safety data from clinical trials of antidepressants have led to a black-box warning from the FDA regarding the use of these medications in children and adolescents and a recommendation for close monitoring. The exact wording of the FDA recommendation is, “all pediatric patients being treated with antidepressants for any indication should be observed closely for clinical worsening, suicidality, and unusual changes in behavior, especially during the initial few months of a course of drug therapy, or at times of dose changes, either increases or decreases.”

Although the frequency of monitoring has been controversial, the FDA further suggested that, “Ideally, such observation would include at least weekly face-to-face contact with patients or their family members or caregivers during the first 4 weeks, then at biweekly visits for the next 4 weeks, then at 12 weeks, and as clinically indicated beyond 12 weeks. Additional contact by telephone may be appropriate between face-to-face visits.” It should be noted, however, that there is no empirical evidence to support the requirement of weekly face-to-face meetings per se for the first 4 weeks after initiating antidepressant treatment. In fact, evidence from large population-based surveys show high reliability of telephone interviews with adolescent subjects for the diagnosis of depression. Although obtaining a diagnosis is not the same as the elicitation of adverse events while in treatment, this evidence suggests that telephone contact may be just as effective in monitoring for adverse events. A regular and frequent monitoring schedule should be developed, and care should be taken to obtain input from the adolescents and families to ensure compliance with the monitoring strategy. Recommendations endorsed by the AACAP have also highlighted the lack of research evidence to support weekly face-to-face visits. However, the AACAP does recommend that providers attempt to follow the FDA guidelines until other research findings become available.

### ONGOING MANAGEMENT

**Recommendation 1:** Systematic and regular tracking of goals and outcomes from treatment should be performed, including assessment of depressive symptoms and functioning in several key domains: home, school, and peer settings (grade of evidence: D; strength of recommendation: very strong).

Goals should include both improvement in functioning status and resolution of depressive symptoms. Tracking of goals and outcomes from treatment should include function in several important domains (ie, home, school, peers). Evidence from large RCTs demonstrates that depressive symptoms and functional impairments may not improve at the same rate with treatment. Therefore, symptoms and functioning should be tracked regularly during the course of treatment with information gathered from both the patients and their families when possible.

According to expert consensus, patients should be seen within 1 week of the initiation of treatment. At every visit, clinicians should inquire about ongoing depressive symptoms, risk of suicide, possible adverse effects from treatment (including the use of specific ad-

---

**TABLE 3: SSRI Titration Schedule**

<table>
<thead>
<tr>
<th>Medication</th>
<th>Starting Dose, mg/d</th>
<th>Increments, mg</th>
<th>Effective Dose, mg</th>
<th>Maximum Dose, mg</th>
<th>Contraindication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citalopram</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>60</td>
<td>MAOIs</td>
</tr>
<tr>
<td>Fluoxetine</td>
<td>10</td>
<td>10–20</td>
<td>20</td>
<td>60</td>
<td>MAOIs</td>
</tr>
<tr>
<td>Fluvoxamine</td>
<td>50</td>
<td>50</td>
<td>150</td>
<td>300</td>
<td>MAOIs</td>
</tr>
<tr>
<td>Paroxetine</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>60</td>
<td>MAOIs</td>
</tr>
<tr>
<td>Sertraline</td>
<td>25</td>
<td>12.5–25</td>
<td>50</td>
<td>200</td>
<td>MAOIs</td>
</tr>
<tr>
<td>Escitalopram</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>MAOIs</td>
</tr>
</tbody>
</table>
verse-effect scales), adherence to treatment, and new or ongoing environmental stressors.

Recently, Emslie et al.\(^6\) examined medication maintenance after response. The researchers randomly assigned those pediatric patients who had responded to fluoxetine by 19 weeks to placebo or to medication continuation for an additional 32 weeks. Of the 20 subjects who were randomly assigned to the 32-week medication relapse-prevention arm, 10 were exposed to fluoxetine for 51 weeks. Significantly fewer relapses occurred in the group of subjects who were randomly assigned to medication maintenance, which suggests that longer medication-continuation periods, possibly 1 year, may be necessary for relapse prevention. In addition, Emslie et al.\(^6\) found the greatest risk of relapse to be in the first 8 to 12 weeks after discontinuing medication, which suggests that after stopping an antidepressant, close follow-up should be encouraged for at least 2 to 3 months.

With the limited evidence in children and adolescents and the emerging evidence in the adult literature that suggests antidepressant medication should be continued for 1 year, the GLAD-PC and AACAP experts concluded that medication should be maintained for 6 to 12 months after the full resolution of depressive symptoms.\(^1,7\)

However, regardless of the length of treatment, all patients should be monitored on a monthly basis for 6 to 12 months after the full resolution of symptoms.\(^19\) If the depressive episode is a recurrence, clinicians are encouraged to monitor patients for up to 2 years given the high rates of recurrence as demonstrated in the adult literature, in which maintenance treatment in those with recurrent depression continue for up to 2 years after the full resolution of symptoms. Clinicians should obtain consultation from mental health if a teen develops psychosis, suicidal or homicidal ideation, or new or worsening of comorbid conditions.

**Recommendation 2:** Diagnosis and initial treatment should be reassessed if no improvement is noted after 6 to 8 weeks of treatment (grade of evidence: B; strength of recommendation: very strong). Mental health consultation should be considered (grade of evidence: D; strength of recommendation: very strong).

If improvement is not seen within 6 to 8 weeks of treatment, mental health consultation should be considered. Evidence of improvement may include reduction in the number of depressive symptoms, improved functioning in social or school settings, or improvement spontaneously reported by the adolescent and/or parent/caregiver. The clinician should also reassess the initial diagnosis, choice and adequacy of initial treatment, adherence to treatment plan, presence of comorbid conditions (eg, substance abuse) or bipolar symptoms that may influence treatment effectiveness, and new external stressors. If a patient has no response to maximum therapeutic dose of an antidepressant medication, the clinician should consider changing the medication. Alternatively, if the patient’s condition fails to improve on antidepressant medication or therapy alone, the addition of, or switch to, the other modality should be considered.

**Recommendation 3:** For patients who achieve only partial improvement after PC diagnostic and therapeutic approaches have been exhausted (including exploration of poor adherence, comorbid disorders, and ongoing conflicts or abuse), a mental health consultation should be considered (grade of evidence: D; strength of recommendation: very strong).

If a patient only partially improves with treatment, mental health consultation should be considered. The clinician should also review the diagnosis and explore possible causes of partial response such as poor adherence to treatment, comorbid disorders, or ongoing conflicts and/or abuse. These causes may need to be managed first before changes to the treatment plan are made.

If a patient has been treated with an SSRI (maximum tolerated dosage) and has shown only partial improvement, the addition of an evidence-based psychotherapy should be considered if it has not previously been conducted. Other considerations may include the addition of another medication, an increase of the dosage above FDA-approved ranges, or a switch to another medication, preferably in consultation with a mental health professional. Likewise, if a patient’s condition fails to improve after a trial of either CBT or IPT and has not yet begun medication, the clinician should consider a trial of SSRI antidepressant treatment. Strong consideration should also be given to a referral to mental health services.

**Recommendation 4:** PC clinicians should actively support depressed adolescents who are referred to mental health to ensure adequate management (grade of evidence: D; strength of recommendation: very strong). PC clinicians may also consider sharing care with mental health agencies/professionals when possible (grade of evidence: B; strength of recommendation: very strong). Appropriate roles and responsibilities regarding the provision and coordination of care should be communicated and agreed upon by the PC clinician and the mental health specialist (grade of evidence: D; strength of recommendation: very strong).

PC clinicians should continue follow-up with adolescents with depression who have been referred to mental health services for assessment and/or management. When possible, PC clinicians may consider sharing management of depressed adolescents with mental health agencies/professionals. There is emerging evidence from the adult literature about the greater effectiveness of “shared-care” models for the management of depression in the PC setting.\(^67,72,79,81\) Similar evidence from case reports in the pediatric literature is emerging.\(^82\)
DISCUSSION
The recommendations regarding treatment and ongoing management highlight the need for PC professionals to become familiar with the use of empirically tested treatments for adolescent depression including both antidepressants and psychotherapy. In particular, antidepressant treatments can be useful in certain clinical situations in the PC setting. However, in many of these clinical scenarios, PCPs need to ensure that there is systematic and regular follow-up and adequate mental health support if needed. The need for systematic follow-up, whether by the PCP or by a mental health provider, is especially important in light of the recent FDA warnings regarding the emergence of adverse events with antidepressant treatment.

Psychotherapy is also recommended as first-line treatment for depressed adolescents in the PC setting. Although the provision of psychotherapy may be less feasible and practical within the constraints (ie, time, availability of trained staff) of PC settings, there is some evidence that quality improvement projects that involve psychotherapy can improve the care of depressed adolescents.27

Another critical recommendation of the guidelines is the need for PCPs to establish connections to available mental health resources in the community, because PCPs will undoubtedly encounter complex cases in which mental health consultation or shared care may be required. Furthermore, increased coordination of care involving different providers are linked to improved outcomes for youth with both general medical and mental health disorders.35,83–86 However, to increase linkages between PCPs and mental health specialists, changes in many existing health care systems need to occur (eg, mental health specialists to set aside time and be reimbursed for brief telephone consultations to PCPs).

The GLAD-PC was developed on the basis of the needs of PC clinicians who are faced with the challenge of caring for depressed adolescents and encounter many barriers including the shortage of mental health resources in most community settings. Although it is clear that more evidence and research in this area are needed, these guidelines represent a necessary step toward improving the care of depressed adolescents in the PC setting. Similar guidelines have also been produced for other health care contexts such as in the United Kingdom (www.nice.org.uk/pdf/CG028NICEguideline.pdf). The GLAD-PC and the toolkit reflects the coming together of available evidence and the consensus of a large number of experts representing a broad spectrum of specialties and advocacy organizations within the North American health care context. However, no improvements in care will be achieved if changes do not occur in the health care systems that would allow for increased training in mental health for PC clinicians and in collaborative models for both PC and specialty care clinicians.

Therefore, it is critical that training programs for PCPs increase their focus on mental health issues and that trainees in both PC and specialty care areas be helped to hone their skills in working in collaborative care models. For providers who are currently practicing, continuing education for primary and specialty care professionals must strengthen skills in collaborative work, and specifically, for PCPs, increase skills and knowledge in the management of depression.

Limitations
Although these guidelines cover a range of issues regarding the management of adolescent depression in the PC setting, there were other controversial areas that were not addressed in these recommendations. These included such issues as universal screening, using a second antidepressant when patients’ conditions fail to respond to an initial antidepressant, and the treatment of subthreshold symptoms. New emerging evidence may impact on the inclusion of such areas in future iterations of the guidelines and the accompanying toolkit. Many of these recommendations are made in the face of absence of evidence or lower levels of evidence.

Future Directions
Ample evidence exists to indicate that guidelines alone are insufficient in closing the gaps between recommended versus actual practices.87,88 Thus, it will be necessary to identify effective methods for disseminating information and to provide assistance in changing practice to PC clinicians. Future studies of these guidelines must build on this work by piloting and evaluating methods, tools, and strategies to facilitate the adoption of these guidelines for the management of adolescent depression in PC settings. These studies must also explore optimal methods for helping clinicians and their organizations/practices address the range of obstacles that may interfere with adoption of necessary practices to yield sustainable management of adolescent depression in PC settings. Also, of course, such studies must show not only changes in PC clinicians’ adolescent depression management but also improvements in outcomes of youth with depression.

Many jurisdictions have recognized the need to increase collaborative care to address the care of adolescents with mental illness. In Canada and the United States, models of care that involve mental health and PC are being implemented. However, the empirical support for these models is modest; therefore, additional research is urgently needed. Work has already begun in Massachusetts to implement GLAD-PC in pediatric practices with funding from AACAP. The findings from this and other studies will build the empirical base for new models of care in the pediatric setting.
**APPENDIX: PART II TOOLKIT ITEMS**

- Algorithm/flow sheet (Fig 1)
- Treatment choices: active support guide, psychotherapy guide, and medication guide and dosage charts
- Referral information
- Authorization to disclose protected health information between PCP and mental health professionals
- Follow-up scripts for management
- Fact sheet/family education materials
- Self-management tools

**ACKNOWLEDGMENTS**

We thank the Center for Substance Abuse Treatment (Substance Abuse and Mental Health Services Administration), the Josiah Macy, Jr Foundation, the New York State Office of Mental Health, the Lowenstein Foundation, the Center for the Advancement of Children’s Mental Health (Columbia University), Sunnybrook Health Sciences Centre (University of Toronto), the American Academy of Pediatrics (District II, New York chapters 1, 2 and 3), the New York Council on Child and Adolescent Psychiatry, the Children’s Health Forum (New York Academy of Medicine), the Kellogg Foundation, and the Civic Research Institute, Inc for financial support of the GLAD-PC project.

The GLAD-PC Steering Group consists of members of the GLAD-PC project team, steering committee, and official organizational liaisons. The GLAD-PC project team members are Peter S. Jensen, MD (project director, REACH Institute), Amy Cheung, MD (project coordinator, University of Toronto/Columbia University), Rachel A. Zuckerbrot, MD (project coordinator, Columbia University), Kareem Ghalib, MD (Columbia University), and Anthony Levitt, MD (project consultant, University of Toronto). The steering committee members are (listed alphabetically) Boris Birmaher, MD (Western Psychiatric Institute & Clinic, University of Pittsburgh), John Campo, MD (Ohio State University and Nationwide Children’s Hospital), Greg Clarke, PhD (Center for Health Research, Kaiser Permanente), Dave Davis, MD (University of Toronto), Angela Diaz, MD (Mount Sinai School of Medicine), Allen Dietrich, MD (Dartmouth Hitchcock Medical Center), Graham Emslie, MD (University of Texas Southwestern Medical School), Bernard Ewigman, MD (Department of Family Medicine, University of Chicago), Eric Fombonne, MD (McGill University), Sherry Gled, PhD (Columbia University), Kimberly Eaton Hoagwood, PhD (Office of Mental Health, New York State/Columbia University), Charles Homer, MD (National Initiative for Children’s Healthcare Quality), Danielle Laraque, MD (AAP New York Chapter 3, District II/Mount Sinai School of Medicine), Miriam Kaufman, MD (Hospital for Sick Children, University of Toronto), Kelly J. Kelleher, MD (Ohio State University), Stanley Kutchera, MD (Dalhousie Medical School), Michael Malus, MD (Department of Family Medicine, McGill University), James Perrin, MD (Massachusetts Medical School/Harvard Medical School), Harold Pincus, MD (Columbia University/New York State Psychiatric Institute), Brenda Reiss-Brennan, APRN (Intermountain Health), Diane Sacks, MD (Canadian Paediatric Society), Ruth E. K. Stein, MD (Forum for Child Health, New York Academy of Medicine, Albert Einstein College of Medicine), and Bruce Waslick, MD, Baystate Health Systems, MA). The organizational liaisons are Angela Diaz, MD (American Academy of Pediatrics), Kelly Kelleher, MD (American Academy of Pediatrics), James Perrin, MD (American Academy of Pediatrics), Diane Sacks, MD (American Academy of Pediatrics/Canadian Paediatric Society), Bruce Waslick, MD (American Medical Association), David Fassler, MD (AACAP), Eric Fombonne, MD (Canadian Academy of Child Psychiatry and Canadian Psychiatric Association), James McIntyre, MD (American Psychiatric Association), Judy Garber, PhD (American Psychological Association), Vicky Wolfe, PhD (Canadian Psychological Association), Michael Malus, MD (College of Family Medicine of Canada), Johanne Renaud, MD (Canadian Association for Adolescent Health), Debbie Ebner, PhD (Society for Adolescent Medicine), Stanford Friedman, MD (Society for Developmental and Behavioral Pediatrics), Terry Stancin, PhD (Society for Developmental and Behavioral Pediatrics), Kathryn Salisbury, PhD (Mental Health Association of New York City), Michael Faenza, MSSW (National Mental Health Association), Susan Bergeson (Depression and Bipolar Support Alliance), Darcy Gruttadaro (National Alliance on Mental Illness), Sandra Spencer (Federation of Families for Children’s Mental Health), and Elizabeth Hawkins-Walsh, DNSc, CPNP (National Association for Pediatric Nurse Practitioners).

In addition to members of the GLAD-PC Steering Group, we also acknowledge the GLAD-PC conference attendees: Perry Adler, PhD; Joan Asarnow, PhD; Lorretta Young Au, MD, FAAP; Abraham Bartell, MD; Tamar Bauer; Rachel Bergeson, MD; Blanche Benenson, MD; Linda Theil Cahill, MD; Wayne Cannon, MD; Marie Barone Casalino, MD; Sonia Chehil, MD; Joseph Cramer, MD; Cathryn Cunningham, MD; Wendy Davis, MD; Carolyn Dewa, PhD; Benard P. Dreyer, MD; M. Flament, MD; D. Clare Fried, MD; William Gardner, Ph.D; Neville Golden, MD; Catherine Goodfellow, MD; Myla Harrison, MD; Sarah Horwitz, PhD; Barbara Huff; Daniel Hyman, MD; Maria Kovacs, PhD; Deborah Launer; Susan Lippert Levitzky, MD; Robert Lubarsky, MD; Christopher P. Lucas, MD; Wanda McCoy, MD; Thomas McInerny, MD; Jessica Mass-Levitt, Ph.D; Margaret McHugh, MD; Laura Mufson, PhD; Gwen Nilsson, MD; Elizabeth Pappadopulos, PhD; Matthew Perkins, MD; Ellen Perrin, MD; Kathryn Salisbury; Marcie Beth.
REFERENCES

17. Jensen PS. Closing the evidence-based treatment gap for children’s mental health services: what we know versus what we do. Emotional Behav Disord Youth. 2002;2:43–50
34. Straw BA, Friedman R, eds. A System of Care for Children and Youth With Severe Emotional Disturbances. Washington, DC: Georgetown University Child Development Center, CASSP, Technical Assistance Center; 1986


Newborn Screening for Pompe Disease: Synthesis of the Evidence and Development of Screening Recommendations

Alex R. Kemper, MD, MPH, MS\textsuperscript{a}, Wuh-Liang Hwu, MD\textsuperscript{b,c}, Michele Lloyd-Puryear, MD, PhD\textsuperscript{d}, Priya S. Kishnani, MD\textsuperscript{e}

\textsuperscript{a}Program on Pediatric Health Services Research and \textsuperscript{b}Division of Medical Genetics, Department of Pediatrics, Duke University, Durham, North Carolina; \textsuperscript{c}Department of Pediatrics, National Taiwan University Hospital, Taipei, Taiwan; \textsuperscript{d}Department of Pediatrics, National Taiwan University College of Medicine, Taipei, Taiwan; \textsuperscript{e}Genetic Services Branch, Division of Services for Children With Special Health Care Needs, Maternal and Child Health Bureau, Rockville, Maryland

ABSTRACT

BACKGROUND. Pompe disease is a lysosomal storage disorder that leads to the accumulation of glycogen and subsequently to muscle weakness, organ damage, and death. Pompe disease is detectable through newborn screening, and treatment has become available recently.

OBJECTIVE. Our goal was to review systematically all available evidence regarding screening for infantile Pompe disease to help policy makers determine whether Pompe disease should be added to their state’s newborn screening battery.

METHODS. We searched online databases, including Medline, clinicaltrials.gov, and the Computer Retrieval of Information on Scientific Projects database, as well as Web sites maintained by federal organizations (eg, the Food and Drug Administration) and other nonprofit or private organizations (eg, the March of Dimes and Genzyme Corp), by using the terms “glycogen storage disease type II,” “Pompe disease,” and “Pompe’s disease.” We also obtained preliminary findings from a screening program in Taiwan. Data were critically appraised and extracted by 2 investigators, one who is an expert in systematic review methods and the other who is an expert in Pompe disease.

RESULTS. The prevalence of Pompe disease has been estimated to be ~1 case per 40 000. Small studies suggest that enzyme therapy is highly efficacious in infantile Pompe disease and that earlier intervention leads to improved outcomes. Screening cannot distinguish between infantile and late-onset Pompe disease. The current screening program in Taiwan has a high false-positive rate; however, the threshold was purposely set low to ensure that no case would be missed.

CONCLUSIONS. Pilot studies of screening are needed to identify the most efficacious strategy for screening and determine how to manage cases of late-onset Pompe disease before screening for Pompe disease is adopted widely by newborn screening programs.
Pompe disease is an autosomal recessive disorder that leads to a deficiency of the enzyme acid α-glucosidase (GAA), resulting in the accumulation of lysosomal glycogen. All patients with Pompe disease share the same underlying lysosomal enzyme deficiency, and all patients have a steady accumulation of glycogen substrate, leading to progressive muscle damage and organ failure; however, the rates of substrate accumulation and tissue damage are variable, reflecting the amount of active enzyme that affected individuals produce. There is also variability in disease progression because of other, less-well understood factors, including modifier genes and the environment.

The American College of Medical Genetics classifies the condition into 2 broad categories: infantile and late-onset disease. The infantile category includes the classic infantile form of the disease, which is the most severe end of the disease spectrum. Individuals with classic infantile Pompe disease usually die in the first year of life, as a result of profound hypotonia and hypertrophic cardiomyopathy. The infantile variant, also referred to as the nonclassic infantile form, has a slower progression to morbidity and death than the classic infantile form. The late-onset category includes childhood and adult-onset forms of Pompe disease. As with infantile Pompe disease, late-onset Pompe disease is associated with progressive muscle weakness and death attributable to respiratory failure. However, hypertrophic cardiomyopathy is not typically associated with the late-onset form of Pompe disease. Age of onset does not always delineate subtypes well; so-called juvenile-onset or mild variant cases occasionally present before 12 months of age. Therefore, the clinical presentation must be considered with the age of onset in the classification of cases.

In 2006, the US Food and Drug Administration (FDA) approved licensure for Myozyme (alglucosidase alfa [Genzyme Corp., Cambridge, MA]), the first effective treatment for Pompe disease. The drug has been approved by the European Union (European Medicines Agency) as therapy for Pompe disease. Because infantile Pompe disease is uniformly lethal and screening is possible with dried blood spots, there has been significant interest in adding Pompe disease to newborn screening programs. This report thus differs from other reviews designed to guide policy decisions, such as those produced by the US Preventive Services Task Force.

METHODS

Primary and Supplemental Data Sources
We searched Medline for all studies published in English from 1966 through July 2006, using the National Library of Medicine Medical Subject Headings term “glycogen storage disease type II” and the keywords “Pompe disease” and “Pompe’s disease,” which identified 614 publications. To ensure completeness of the literature search, we reviewed reference lists and articles from the authors’ libraries. Although no additional published articles were identified in these libraries, we did identify 1 study in press. We supplemented the primary literature search by searching the Web sites of the American College of Medical Genetics, the American Academy of Pediatrics, the National Newborn Screening Resource Center, the March of Dimes, the Acid Maltase Deficiency Association, the International Pompe Association, the Computer Retrieval of Information on Scientific Projects database of research projects and programs supported by the Department of Health and Human Services, a database of federally and privately supported clinical research (clinicaltrials.gov), the FDA, and Genzyme Corp. We also obtained preliminary data from a pilot project on screening newborns in selected Taiwanese hospitals.

Data Extraction

We extracted data to answer key questions regarding the natural history and burden of suffering related to infantile Pompe disease, methods for screening and diagnosis, effectiveness of treatment, screening and treatment recommendations, and ongoing research. Although our focus was on infantile Pompe disease, we recognize that screening may detect late-onset Pompe disease; therefore, we also searched for data regarding the number of infants at risk for late-onset Pompe disease who would be detected and the benefit of intervention for asymptomatic infants at risk for late-onset Pompe disease. Because Pompe disease is a rare condition, we chose to consider a wide range of study designs, instead of restricting data on the basis of the quality of the study design (eg, randomized, controlled trials or other prospective designs). We considered case reports, case series, uncontrolled intervention trials, and expert consensus. However, we did exclude research studies that did not include human subjects. Two reviewers, one an expert in systematic reviews and policy issues related to newborn screening (Dr Kemper) and the other an expert in Pompe disease (Dr Kishnani), worked together to synthesize the data for this report. There was an insufficient number of studies to allow for meaningful meta-analysis.
RESULTS

What Is the Natural History of Infantile Pompe Disease and Its Burden of Suffering?

An international retrospective study based on chart reviews for children with infantile (classic or nonclassic) Pompe disease found that the median age at which children first became symptomatic was 2 months and the median age of diagnosis was ~5 months.3 The median age of death was ~9 months, with 9% of patients still alive at 24 months of age. Longer survival times were associated with higher skin fibroblast GAA activity. Similar morbidity results were found in an earlier retrospective cohort study.6 A review of the experience at 1 large medical institution identified 30 infants with Pompe disease.7 The average ages of diagnosis and death in that cohort were 5.1 month and 8.6 months, respectively. None of those infants survived past 12.4 months. Infants who survive into their second year of life most likely have the nonclassic form of the disease.8

The incidence of Pompe disease, on the basis of diagnosed cases, has been estimated to range from 0.17 to 1.31 cases per 100 000.3-11 However, these estimates are highly dependent on disease recognition; therefore, the true incidence could be underestimated because of case ascertainment bias. Other studies have attempted to estimate the incidence on the basis of extrapolation from carrier frequency.

The gene for GAA is located on chromosome 17.12 Many different mutations have been described, and some specific mutations have been associated with a particular disease course.13 Approximately 3000 anonymous dried blood spots in the Netherlands were evaluated for 3 specific mutations, 2 associated with infantile Pompe disease and 1 associated with late-onset Pompe disease. These mutations represented 63% of those associated with Pompe disease among the Dutch.14 On the basis of the carrier frequency of these mutations (infantile form: 1 case per 284; late-onset form: 1 case per 154), the incidence of infantile Pompe disorder associated with these mutations only was estimated to be ~1 case per 138 000 (95% confidence interval: 1 case per 43 000 to 1 case per 536 000) and that of late-onset disease to be 1 case per 57 000 (95% confidence interval: 1 case per 28 000 to 1 case per 128 000). An earlier study estimated the total incidence of Pompe disease, regardless of type, to be ~1 case per 40 000, on the basis of carrier testing of randomly selected individuals from New York for 7 mutations that were thought to be responsible for 29% of the cases of Pompe disease.15 No published data are available regarding the economic burden of infantile Pompe disease or the impact of the disease on families.

What Treatments Are Available for Pompe Disease?

Before the approval of alglucosidase alfa for enzyme replacement therapy, only supportive treatment was available. Although aggressive supportive care can improve nutrition and temporarily improve strength, the risk of morbidity is not altered.23 Oral supplementation with l-alanine or a diet high in branched-chain amino acids may improve the myopathy associated with nonclassic infantile Pompe disease by reducing protein turnover and resting energy expenditure.24-26 Bone marrow transplantation has not been effective, perhaps because it has not been performed sufficiently early in the disease course or because children with classic infantile Pompe disease die before engraftment can occur.27-29

Recombinant human GAA (rhGAA) for human clinical trials has been produced from rabbit milk30 and from Chinese hamster ovary cells. Alglucosidase alfa, the FDA-approved rhGAA form, is produced from Chinese hamster ovary cells. The recommended dose of alglucosidase alfa is 20 mg/kg, administered intravenously ev-
ery 2 weeks in ~4 hours. The current average wholesale price for the treatment is $720 per 50-mg vial.31

The first published phase I/II trial of rhGAA derived from Chinese hamster ovary cells included 3 infants (2.5, 3, and 4 months of age) with infantile Pompe disease who were given rhGAA (5 mg/kg) twice weekly.7 All patients survived beyond 1 year of life. After 1 year of therapy, 2 of the children had decreases in heart size related to improvement in cardiomyopathy, and the third maintained normal heart status despite hypertrophy. Two of the patients developed high and sustained anti-rhGAA antibody titers, which were associated with declines in motor development and pulmonary function. Those 2 patients became ventilator-dependent. However, the third patient, who is now almost 8 years of age, is ambulatory and does not require ventilator support or supplemental oxygen therapy (P.S.K., unpublished data).

A subsequent study of rhGAA from Chinese hamster ovary cells enrolled 8 children with classic infantile Pompe disease that was diagnosed between 1.8 and 6.5 months of age.32 Treatment with rhGAA began between 2.7 and 14.6 months of age. Six of the children survived a 52-week treatment period and 5 showed motor improvement, including 3 who gained the ability to walk. During the 52-week treatment period and a subsequent extension phase, a total of 6 of the 8 patients died, with a median age of death of 21.7 months. The 2 surviving children, both of whom began treatment at 6 months of age, are now >4 years of age.

A phase II trial of rhGAA derived from rabbit milk enrolled 2 infants (3.1 and 5.9 months of age) with classic Pompe disease.33 After 48 weeks, both children experienced improvements in motor development and cardiac function and neither required mechanical ventilator support or supplemental oxygen therapy (P.S.K., unpublished data).

Does Treatment in the Presymptomatic or Early Symptomatic Phase of Pompe Disease Lead to Better Health Outcomes?

No data are available regarding treatment for presymptomatic children with Pompe disease. Data from the screening pilot study in Taiwan should provide this information. However, earlier treatment seems to confer benefit. In a trial with 18 children with classic infant Pompe disorder who began treatment by 26 weeks of age,4 the children seemed to have better survival rates and improved motor outcomes, compared with the children in the previously described studies in which children began therapy at an older age. In that study, all children survived the 52-week treatment period, 3 patients required invasive ventilatory support, none had cardiac failure, 7 could walk, 3 could stand independently, and 3 could sit independently.4

What Is the Potential Harm of Screening for Pompe Disease?

On the basis of the previously presented epidemiologic data, screening for Pompe disease may identify ≥2 cases of late-onset Pompe disease for each case of infantile Pompe disease. No data are available regarding the management of early-detected or presymptomatic late-onset disease, including when to begin enzyme replacement therapy. No data are available regarding the benefit or harm of such early detection on families. Although research on fragile X syndrome suggests that parents regard presymptomatic diagnosis as valuable,42 specific research is needed for Pompe disease.
As with any screening program, there would also be false-positive results. No data are available regarding the impact of these specific false-positive results on newborns and their families; false-positive newborn screening results have been associated with increased family stress and parent/child dysfunction.43

**What Has Been the Experience of Screening Programs for Pompe Disease?**

A Pompe disease newborn screening pilot program began in Taiwan in October 2005, using a fluorometric assay (W.-L.H., unpublished data). The program was designed to be highly sensitive, to ensure that no cases would be missed, although the false-positive rate would necessarily increase as a result. Complete details of this screening program will be presented elsewhere. However, as of July 2006, ~0.9% of the nearly 71,000 children who were screened were recalled for a second blood sample. Of those children, ~9% (~<0.1% of the initial cohort) had positive second-tier test results, from which 3 cases of classic infantile Pompe disease and no cases of late-onset disease were detected. In Taiwan, dried blood spots for newborn screening are collected on the third day of life. Earlier screening is not thought to affect test accuracy.

**What Are the Recommendations of Professional Organizations Regarding Screening for Pompe Disease?**

Professional organizations, including the American Academy of Pediatrics and the American Academy of Family Physicians, advocacy organizations, including the March of Dimes and the Genetic Alliance, and the Department of Health and Human Services Advisory Committee on Heritable Disorders and Genetic Diseases in Newborns and Children support the recommendations of the American College of Medical Genetics Newborn Screening Expert Group.44 The Newborn Screening Expert Group did not recommend screening for Pompe disease, primarily because neither enzyme replacement therapy nor a validated screening test was available at the time of its review.

**What Ongoing Research Would Help Answer Questions About the Value of Newborn Screening for Pompe Disease?**

Genzyme Corp, which manufactures alglucosidase alfa, continues to evaluate the safety and efficacy of the therapy. In addition, Genzyme maintains a registry of individuals diagnosed as having Pompe disease.45 This registry excludes individuals enrolled in a Genzyme-sponsored clinical trial. As of September 2006, there were 35 children with infantile Pompe disease in the registry (C. Dandrea, MBA, written communication, 2006). In addition, the International Pompe Association and the Erasmus Medical Center (Rotterdam, Netherlands) maintain a registry of individuals with late-onset Pompe disease. Findings from secondary analyses of data from these registries could lead to a better understanding of the course of treated Pompe disease and identify strategies to improve care. These registries may also be important for health system planning, by identifying the resources and services needed to provide care for those with Pompe disease. However, we are unaware of any ongoing health services research focusing on Pompe disease, including modeling of the potential impact on costs and benefits of newborn screening for Pompe disease.

There is significant ongoing basic science and early translational research in Pompe disease. For example, the National Institutes of Health are funding research into the pathogenesis (eg, Genetic Metabolic Myopathies [principal investigator: Dr Paul Plotz]) and the development of novel therapies, such as gene therapy (eg, “Correction of Inherited Cardiomyopathy Using Adeno-associated Virus Vectors” [principal investigator: Dr Barry Byrne] and “Gene Delivery to Striated Muscle by Systematic Adeno-associated Virus Vectors” [principal investigator: Dr Dwight Koeberl]).

**DISCUSSION**

Determination of whether to recommend that newborn screening should include a new condition, such as Pompe disease, is complex and involves assessment of the potential benefits and harm of screening from the perspectives of families and society.46 Although research into management is still underway, infantile Pompe disease is a uniformly fatal condition; enzyme replacement therapy improves both length and quality of life for affected children. However, there are still important unanswered questions about the harm of screening.

Current testing strategies identify those who have late-onset Pompe disease. This problem is not unique to Pompe disease. For example, individuals with short-chain acyl-coenzyme A dehydrogenase deficiency can have highly variable clinical outcomes, which cannot be predicted through identification in the newborn period.47 However, most states detect short-chain acyl-coenzyme A dehydrogenase in newborn screening.47 We think that parents who elect to have their children tested for Pompe disease should be informed that there is the possibility of detecting late-onset Pompe disease and that there is little knowledge about how to provide care for asymptomatic children with late-onset Pompe disease. It is important to recognize that no studies have been performed to determine how to communicate effectively this complex information about potential harm, including family anxiety, medicalization of a condition that may not cause problems for many years, and long-term consequences related to insurability and employability.

The recent data from the pilot project of screening in Taiwan showed a high rate of false-positive results. A high rate of false-positive results would overwhelm the ability of public health programs in the United States to ensure appropriate follow-up evaluation. Before screen-
ing for Pompe disease can be recommended to be added to states’ newborn screening panels, pilot studies are needed to develop testing strategies to ensure that testing is adequately sensitive, while greatly decreasing the number of false-positive results.

No data are available regarding whether there is sufficient infrastructure to care for children who would require diagnostic confirmation or treatment. Without the availability of appropriate services first being ensured, some children and their families may also suffer harm.

Finally, no data are available regarding the costs associated with screening and treatment for Pompe disease. Such costs are not the only determinant of whether to initiate screening; however, it is important for policy makers to understand explicitly the trade-offs associated with any new public health initiative.

Most of the evidence gathered in this report is available in the peer-reviewed literature, but this review would be incomplete without the inclusion of data obtained through personal communication. We recognize that unpublished data may be less reliable, because they are not subjected to the important scrutiny of peer review. We also recognize that there may be important unpublished data of which we are unaware. Developing methods to incorporate non–peer-reviewed observations into systematic reviews that are both transparent and replicable will be central to developing reports to guide screening policy for the many rare conditions for which screening is becoming available.

On the basis of our experience in preparing this report, we also think that it would assist policy makers if a standardized recommendation scheme were developed. The system used by the US Preventive Services Task Force is a good model; screening recommendations are graded from A (“good evidence that [the service] improves important health outcomes and concludes that benefits substantially outweigh harms”) through D (“at least fair evidence that [the service] is ineffective or that harms outweigh benefits”), with I for situations in which “evidence that the [service] is effective is lacking, of poor quality, or conflicting and the balance of benefits and harms cannot be determined.”48 However, because of the lack of high-quality data, we think that many of the new conditions identifiable through newborn screening would receive a grade of I. We propose the following classification system. (1) Universal screening recommended: all newborn screening programs should implement screening for the condition once follow-up infrastructure is in place. (2) Targeted screening recommended: newborn screening programs should target high-prevalence population groups for screening once follow-up infrastructure is in place. An historical example of targeted screening would be testing for sickle cell disease among black children. There is significant concern about the ethics and feasibility of such targeted screening.49 Therefore, we think that the threshold for recommending targeted screening should be high. (3) Pilot study of screening recommended: although there is reasonable evidence to suggest that screening for the condition is likely to be beneficial, there are still important unanswered questions about the impact or performance of screening. Screening (either universal or targeted) has a high likelihood of being recommended once these questions are resolved. (4) Pivotal studies required: although there are many fundamental unanswered questions about the impact of screening, screening on balance may lead to greater benefits than harm. Additional research is needed to answer fundamental questions (eg, natural history and prevalence of the condition, effectiveness of treatment, and accuracy of screening). (5) No general recommendation: the evidence suggests that the potential benefits and harm of screening for the condition are closely balanced. (6) Recommended against: the evidence suggests that screening for the condition would lead to more harm than benefit.

This classification scheme, along with the underlying evidence synthesis, not only could help guide state newborn screening policy makers but also could facilitate coordination of the activities of researchers, screening advocates, private foundations, drug and device manufacturers, and the federal agencies involved in newborn screening (eg, the National Institutes of Health, the Health Resources and Services Administration, the Agency for Healthcare Research and Quality, and the FDA), to identify and to answer the key policy questions regarding newborn screening for any particular condition. Using this classification scheme, we would grade screening for Pompe disease as “pilot study of screening recommended,” because of the clear benefit of alglucosidase alfa for children with infantile Pompe disease. However, significant concerns remain about screening accuracy, the management of identified cases of late-onset Pompe disease, and the challenge of informing parents about the benefits and harm of newborn screening for Pompe disease. Until these concerns have been resolved with scientifically based screening pilot studies, we think that it is too soon to recommend that all states add Pompe disease to their newborn screening panels.

ACKNOWLEDGMENTS
This project was funded by the Maternal and Child Health Bureau of the Health Resources and Services Administration.

REFERENCES


47. Waisbren SE. Newborn screening for metabolic disorders. JAMA. 2006;296:993–995
EXPERIENCE & REASON

Potential Misdiagnosis of 3-Methylcrotonyl-Coenzyme A Carboxylase Deficiency Associated With Absent or Trace Urinary 3-Methylcrotonyglycine

Lynne A. Wolfe, MS, CRNPa,b, David N. Finegold, MDa,b,c,d, Jerry Vockley, MD, PhD,a,b,c, Nicole Walters, RD, LDN,a,b, Celine Chambaz, BS,a, Terttu Suormala, PhD,d, Hans Georg Koch, MD,a, Dietrich Matern, MD,b, Bruce A. Barshop, MD, PhD,d, Lorna J. Cropcho,a,b,d, Matthias R. Baumgartner, MD,a, K. Michael Gibson, PhD,a,b,c,d

aDivision of Medical Genetics, Departments of Pediatrics, Human Genetics, and Pathology; bBiochemical Genetics/Nutrition Laboratory, Children’s Hospital of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; cDivision of Metabolism and Molecular Pediatrics, University Children’s Hospital, Zurich, Switzerland; dMetabolic Unit, University Children’s Hospital, Basel, Switzerland; eDepartment of Pediatrics, Children’s Hospital, Braunschweig, Germany; fDepartments of Laboratory Medicine and Pathology, Pediatric and Adolescent Medicine and Medical Genetics, Mayo Clinic College of Medicine, Rochester, Minnesota; gBiochemical Genetics Laboratory, University of California at San Diego, La Jolla, California

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

We report 2 patients with isolated 3-methylcrotonyl-coenzyme A carboxylase deficiency whose urine was devoid of, or contained only trace, 3-methylcrotonyglycine, the pathognomonic marker for this disorder. The first patient, a girl with trisomy 21, was detected through newborn screening with an elevated 5 carbon hydroxycarnitine species level, and the second patient came to clinical attention at the age of 5 months because of failure to thrive and developmental delay. Investigation of urinary organic acids revealed an elevated 3-hydroxyisovaleric acid level but no demonstrable 3-methylcrotonyglycine in both patients. Enzyme studies in cultured fibroblasts confirmed isolated 3-methylcrotonyl-coenzyme A carboxylase deficiency with residual activities of 5% to 7% and 12% of the median control value, respectively. Incorporation of 14C-isovaleric acid into intact fibroblasts was essentially normal, showing that the overall pathway was at least partially functional and potentially explaining the absence of 3-methylcrotonylglycine in urine. Mutation analysis of the MCCA and MCCB genes revealed that both patients were compound heterozygous for a missense mutation, MCCB-c.1015G→A (p.V339M), and a second mutation that leads to undetectable MCCB messenger (poly A+) RNA. Absent or trace 3-methylcrotonyglycine levels in urine raises the potential for misdiagnosis in the clinical biochemical genetics laboratory based solely on urine organic acid analysis using combined gas chromatography-mass spectrometry.
MCCB (MCCC2; Online Mendelian Inheritance in Man No. 609014), respectively.3

The addition of expanded newborn screening for inborn errors of metabolism using tandem mass spectrometry has demonstrated that 3-MCC deficiency is one of the more commonly detected heritable organic acidurias.6-7 Prevailing data suggest that 3-MCC deficiency is, for the larger part, a biochemical phenotype that only manifests in the presence of a major precipitating event.3-11 In addition, asymptomatic mothers with 3-MCC deficiency deliver newborns with no clinical sequelae who nonetheless manifest increased 5 carbon hydroxycarnitine species (C5-OH carnitine) on newborn screening.12 The seemingly asymptomatic nature of 3-MCC deficiency prompted the German government to omit this disease from their list of target disorders for expanded newborn screening.6

Here we report clinical, biochemical, and molecular findings in 2 patients with isolated 3-MCC deficiency whose urine was essentially devoid of 3-MCG, the pathognomonic marker for this disorder. Our results expand the biochemical and molecular heterogeneity of 3-MCC deficiency, with screening ramifications for clinical biochemical genetics laboratories.

MATERIALS AND METHODS

Urine organic acid and plasma/urine acylcarnitine analyses were performed as described.13,14 Activity of 3-MCC was determined in cultured skin fibroblasts, and isolated venous blood lymphocytes (only in patient 2) and 14C-isovaleryl acid incorporation in fibroblasts were determined as described previously.12,15,16 Mutation analysis of the MCCCA and MCCCB genes was performed as described previously.11 All studies were performed with informed consent and institutional review board approval.

CASE REPORTS

PATIENT 1. The proband is the 46XX+21 offspring of a gravida 6, para 7 mother. Delivery and the newborn period were uneventful, and there is no known consanguinity. Parental ethnicity includes Italian, German, Ukrainian, and Swiss ancestry. Six siblings (3 male, 3 female) ranged in age from 18 months to 24 years. All family members (including the proband’s female fraternal twin) had serum acylcarnitine profiling and urine organic acid analysis completed, neither of which demonstrated clinical or biochemical abnormalities.

Triplicate newborn screening samples were remarkable for an elevated C5-OH acylcarnitine level (1.5–1.9 μmol/L [reference: <1.0 μmol/L]), which led to initiation of a full metabolic workup and clinical evaluation. Plasma acylcarnitine analysis confirmed a consistently elevated C5-OH carnitine level (0.19 μmol/L [reference: <0.03 μmol/L]). Retrospective acylcarnitine analysis of a urine sample collected while the patient was supplemented with L-carnitine revealed a significant overexcretion of free carnitine and acetylcarnitine, whereas the excretion of C5-OH carnitine was normal (0.3 mmol/mol creatinine [reference: <0.52 mmol/mol creatinine]). Conversely, follow-up urine acylcarnitine analysis (after L-carnitine supplementation had been discontinued) revealed elevated excretion of C5-OH carnitine (2.18 mmol/mol creatinine). Urine organic acid analysis revealed an elevated 3-HIVA level as the only finding (~86–244 mmol/mol creatinine [reference: <46 mmol/mol creatinine], n = 3). 3-MCG was undetectable or present in only a trace amount. Results of plasma amino acid analysis were normal.

Characteristic dysmorphic features and other minor signs and symptoms (mild blepharitis, strabismus, perinatal granuloma, and gastroesophageal reflux disease) led to a diagnosis of trisomy 21. At 15 months of age, the patient was also diagnosed with celiac disease because of frequent episodes of diarrhea and acute duodenitis. Significant weight loss after the first year of life was corrected with alterations in elemental formula composition. At the time of this writing, the patient is 18 months old, is receiving physical, occupational, and speech therapy, and is developmentally delayed with mild muscular hypotonia consistent with her trisomy 21 diagnosis. She was started on biotin (10 mg daily) and L-carnitine (50 mg/kg per day initially). As her carnitine levels normalized and muscular hypotonia improved, the carnitine was discontinued at ~15 months of age, before her diagnosis of celiac disease. She remains on empiric biotin supplementation without clear evidence of benefit.

PATIENT 2. This proband is the second child of Turkish parents. Mutation analysis was reported previously (patient 008).17 Her elder brother and parents were healthy, and there is no known consanguinity. The pregnancy, birth, and newborn period were uneventful. The patient came to clinical attention at the age of 5 months. She was admitted to the hospital because of failure to thrive and frequent vomiting that had started at the age of 2 months when she was switched from mother’s milk to a commercial infant formula. Physical examination showed a dystrophic child with a weight of 4280 g (1 kg below the third percentile), length of 60 cm (0.5 cm below the 3rd percentile), and head circumference of 40 cm (3rd percentile). She displayed mild muscular hypotonia and delayed psychomotor development.

Repeat urine organic acid analysis revealed 3-HIVA as the only finding (108–217 mmol/mol creatinine [reference: <46 mmol/mol creatinine], n = 8). 3-MCG was undetectable or present in only a trace amount. Analysis of organic acids in serum and cerebrospinal fluid displayed a clearly elevated 3-HIVA concentration in cerebrospinal fluid (17–29 μmol/L [reference: <3 μmol/L], n = 3), whereas the plasma 3-HIVA level was only slightly
increased (8–12 μmol/L [reference: <5 μmol/L], n = 3). The free plasma carnitine level was slightly reduced (20 μmol/L [reference range: 26–49 μmol/L]), whereas the total plasma carnitine level was within the reference range (41.3 μmol/L [reference range: 38–68 μmol/L]) and the esterified carnitine level was slightly increased (21.3 μmol/L [reference range: 7–19 μmol/L]). Plasma amino acid analysis was normal. The patient was started on biotin (10 mg/day) and a low-protein diet (1.3 mg/kg body weight).

Further clinical evaluation revealed hearing loss (brainstem evoked potentials: right ear, moderate; left ear, severe decrease) but no additional pathologies. Results of computed tomography and MRI scans of the brain as well as an electroencephalogram were normal. Her psychomotor development remained severely retarded. As of this writing, the patient is 12 years old but, unfortunately, has been lost to follow-up.

Enzyme Studies
An assay of carboxylase activities of fibroblasts of both patients revealed isolated MCC deficiency (Table 1). The first enzyme assay performed on fibroblasts of patient 1 revealed deficient MCC activity (14–23 pmol/min per mg of protein [reference range 140–350 pmol/min per mg of protein]). However, the absence of urinary 3-MCG prompted further confirmatory enzyme studies in a second laboratory (Table 1). Low levels of residual MCC activity of 5.3% and 6.7% of the median control value were obtained from fibroblasts of patient 1, which is in contrast to <3% obtained in 9 patients with complete MCC deficiency. In accordance with the presence of residual activity, the incorporation of 14C-isovalerate in intact fibroblasts was clearly higher than in patients with complete deficiency and only slightly below the reference range. Multiple carboxylase deficiency (MCD) was excluded by demonstration of normal propionyl-coenzyme A carboxylase (PCC) activity and by failure to increase the activity of MCC or PCC or incorporation of 14C-isovalerate by addition of high concentrations of biotin (10 μmol/L) to the culture medium (results not shown). The relatively low activity of the control enzyme PCC observed in this particular cell line probably reflects its poor proliferation rate. This also may have influenced the level of residual MCC activity and 14C-isovalerate incorporation.

In patient 2, residual MCC activity in fibroblasts was 12% and in lymphocytes 18% of the median reference values, with clearly normal levels of PCC activity and incorporation of 14C-isovalerate into intact fibroblasts (Table 1). As in patient 1, there was no increase of MCC activity in fibroblasts when grown in the presence of a high concentration of biotin (results not shown).

Mutation Analysis
Initial reverse-transcription polymerase chain reaction (PCR) amplification of MCCB and MCCA complementary DNA from fibroblast messenger (poly A +) RNA (mRNA) followed by sequencing of the entire open reading frame revealed that both patients are homozygous for MCCB-c.1015G→A (p.V339M). However, genomic PCR of the corresponding exon showed that both were only heterozygous for pV339M. Despite sequencing of all MCCB exons and flanking intronic sequences, we were not able to detect any additional sequence alterations. In repeated reverse-transcription PCR experiments, pV339M was clearly homozygous, even after treating the fibroblasts with cycloheximide to suppress nonsense-mediated mRNA decay, which suggests that the steady state of mRNA from the second allele was not detectable, as would be the case

<table>
<thead>
<tr>
<th>Subject</th>
<th>Carboxylases, pmol/min per mg of Protein</th>
<th>14C-isovaleric Acid, nmol/16 h per mg of Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MCC</td>
<td>PCC</td>
</tr>
<tr>
<td>Lymphocytes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient 2</td>
<td>67.6</td>
<td>463</td>
</tr>
<tr>
<td>3 patients with complete deficiency</td>
<td>0, 5.5, 6.2</td>
<td>567, 577, 1010</td>
</tr>
<tr>
<td>109 controls, median (range)</td>
<td>370 (119–824)</td>
<td>563 (228–954)</td>
</tr>
<tr>
<td>Fibroblasts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient 1</td>
<td>16.8, 21.5</td>
<td>233, 269</td>
</tr>
<tr>
<td>Patient 2</td>
<td>37.3</td>
<td>695</td>
</tr>
<tr>
<td>9 patients with complete deficiency, median (range)b</td>
<td>1.6 (0–7.0)</td>
<td>906 (305–1198)</td>
</tr>
<tr>
<td>Control values</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>319 (125–831)</td>
<td>740 (207–2150)</td>
</tr>
<tr>
<td>n</td>
<td>73</td>
<td>73</td>
</tr>
</tbody>
</table>

— indicates not applicable.

a Individual values in the patients are the mean of duplicate or triplicate determinations within 1 experiment

b Absence of residual 3-MCC activity in fibroblasts of these patients was confirmed by showing no increase of 3-MCC activity when assayed with increasing incubation time, as demonstrated in Fig 1.16.
for a promotor mutation or an intragenic deletion or insertion undetected by genomic PCR.

DISCUSSION

Expansion of newborn screening programs by use of tandem mass spectrometry has revealed a high incidence of 3-MCC deficiency. This finding has highlighted an ongoing conundrum associated with comprehensive newborn screening, which is the identification of conditions such as 3-MCC deficiency that are largely asymptomatic with minimal penetrance and mild (but variable) clinical expression. Whereas severe phenotypes have been reported in 3-MCC deficiency, most literature reports suggest a benign clinical course, as shown for patient 1. Mild cases of 3-MCC may not require any treatment, and overaggressive intervention may do more harm than good. As data accumulate, then, it remains possible that a low benefit/risk ratio for 3-MCC deficiency may prompt some newborn screening programs in the future to consider omission of this disorder from their target list of disorders.

However, although it would be possible to set the newborn screening system to ignore an elevation in C5-OH acylcarnitine levels, C5-OH acylcarnitine is not unique to 3-MCC deficiency in the newborn or mother. Other conditions, including 3-hydroxy-3-methylglutaryl-coenzyme A lyase (HMG-CoA lyase) deficiency, β-ketothiolase (mitochondrial 2-methylacetoacetyl-coenzyme A thiolase) deficiency, MCD resulting from homocarnitine was excluded from screening programs. Of these 6 conditions, the American College of Medical Genetics included 3-MCC, HMG-CoA lyase deficiency, β-ketothiolase deficiency, and MCD among the 29 core conditions for which each newborn should be screened. 2-Methyl 3-hydroxybutyric acidemia and 3-methylglutaconic aciduria (3-MGA) must be considered in the differential diagnosis of isolated elevation of C5-OH acylcarnitine levels and would be undetected if C5-OH carnitine was excluded from screening programs. Of these 6 conditions, the American College of Medical Genetics included 3-MCC, HMG-CoA lyase deficiency, β-ketothiolase deficiency, and MCD among the 29 core conditions for which each newborn should be screened. 2-Methyl 3-hydroxybutyric acidemia and 3-MGA are considered secondary targets defined as conditions that are included in the differential diagnosis of a primary target disorder but lack the availability of effective treatment strategies. From a diagnostic perspective, in our 2 patients 3-HIVA was clearly elevated, which encompasses 3-MCC deficiency, HMG-CoA lyase deficiency, MCD, some forms of 3-MGA, marked ketosis, and intervention with valproate in the differential diagnosis. All (except 3-MCC deficiency) were excluded by urine organic acid analysis and/or clinical history.

Although novel, the association of 3-MCC deficiency with trisomy 21 in patient 1 is likely a coincidence. The occurrence of both, however, complicates the association of any clinical feature with a single genetic etiology. In patient 2, the long-term neurologic outcome is very unsatisfactory. It is highly likely that this patient’s condition would have been diagnosed in the extended newborn screening programs. Whether early diagnosis and a timely start of therapy and preventive measures would have changed the clinical course remains unknown. Current clinical, biochemical, and genetic data strongly support the idea that factors other than the genotype at the MCC loci must have a major influence on the phenotype of MCC deficiency. Such additional factors that could influence the phenotypic consequences of 3-MCC deficiency include modifying genes and, perhaps more importantly, the extent to which the pathway is stressed by dietary or other environmental factors such as excessive protein breakdown associated with concurrent infection. In patient 2, excessive protein breakdown caused by severe failure to thrive may have led to the severe phenotypic expression of the disease.

An absence or trace of 3-MCG in urine expands the biochemical heterogeneity associated with isolated 3-MCC deficiency. Marked elevation of urine 3-MCG levels has hitherto been considered the pathognomonic marker for 3-MCC deficiency. Before implementation of expanded newborn screening, one wonders how many patients escaped detection by routine gas chromatography/mass spectroscopy analysis of urine organic acids because of the absence of 3-MCG. Koeberl et al presented a single case in which the urinary excretion of 3-MCG in an enzyme-confirmed patient was noted as only trace. As previously demonstrated by Kølvraa and Gregersen, liver acyl-CoA:glycine-N-acetyltransferase is the likely candidate for catalysis of 3-methylcrotonyl-CoA conjugation with glycine (Fig 1). It remains possible that our patients have an abnormality of acetyltransferase activity, located on chromosome 11, in addition to the abnormalities of trisomy 21 in patient 1 and of 3-MCC (chromosomes 3 and 5 for the different subunits) in both patients. Data to support these hypotheses, however, are lacking. Alternatively, it may be possible that the absence of strikingly abnormal organic acid profiles is consistent with a milder 3-MCC phenotype, which would be in accordance with residual 3-MCC activity found in both patients. However, the rather severe phenotypic expression in patient 2 and earlier reports of severely affected patients with residual activity do not support this hypothesis. Despite low levels of 3-MCC activity measured in fibroblasts, incorporation of [14C]-isovalerate in intact cells was virtually normal in both patients, which indicates that 3-MCC is at least partially functional in vitro. In addition, functional expression studies confirm that MCCC-p.V339M confers ~4% residual activity when transfected into MCCC-null cells. Finally, the level of residual activity may vary between different tissues, as illustrated by higher activity in lymphocytes compared with fibroblasts in patient 2 and in 2 other patients reported previously. These findings suggest that the overall leucine-degradative pathway may be
CONCLUSIONS
We have presented 2 probands with isolated 3-MCC deficiency whose urine organic acid profile was devoid or contained only traces of the characteristic metabolite, 3-MCG. The characterization of our patients suggests that the term “3-methylcrotonyglycinuria” may not be used interchangeably in the future to indicate 3-MCC deficiency. In those instances in which acylcarnitine analysis is unavailable for confirmation, the absence of 3-MCG may hinder the correct diagnosis or result in misdiagnosis on the sole basis of an increase in the urinary 3-HIVA level, which represents important information for clinical biochemical genetics laboratories that are involved in detection and follow-up of heritable organic acidurias. This observation underscores the utility of the diagnostic algorithm developed by the American College of Medical Genetics (available at www.acmg.net/resources/policies/NBS/NBS-sections.htm) for the follow-up of elevated C5-OH acylcarnitine concentrations in newborn screening blood spots.

ACKNOWLEDGMENTS
We gratefully acknowledge the Center for Inherited Disorders of Energy Metabolism (Rainbow Babies and Children’s Hospital, Cleveland, OH) for enzymatic studies.

REFERENCES
15. Suormala T, Wick H, Bonjour JP, Baumgartner ER. Rapid differential diagnosis of carboxylase deficiencies and evalua-


Hematopoietic Stem Cell Transplantation Corrects the Immunologic Abnormalities Associated With Immunodeficiency–Centromeric Instability–Facial Dysmorphism Syndrome

Andrew R. Gennery, MDa, Mary A. Slatter, MDa, Robbert G. Bredius, PhDb, Melanie M. Hagleitner, PhDb, Corry Weemaes, PhDb, Andrew J. Cant, MDa, Arjan C. Lankester, PhDb

aPaediatric Immunology Department, Newcastle General Hospital, Newcastle Upon Tyne, United Kingdom; bDepartment of Pediatrics, Bone Marrow Transplant Unit, Leiden University Medical Center, Leiden, Netherlands; cDepartment of Pediatrics, Nijmegen University Medical Center, Nijmegen, Netherlands

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

Immunodeficiency–centromeric instability–facial dysmorphism syndrome, characterized by variable immunodeficiency, centromeric instability, and facial anomalies caused by epigenetic dysregulation resulting in hypomethylation, is caused in many patients by mutations in \( \text{DNMT3B} \), a DNA methyltransferase gene; associated infections are a major cause of serious sequelae and death. Hematopoietic stem cell transplantation may improve the clinical course in immunodeficiency–centromeric instability–facial dysmorphism syndrome. We report 3 unrelated patients with persistent infections and intestinal complications who successfully underwent hematopoietic stem cell transplantation after nonmyeloablative or myeloablative conditioning regimens using HLA-matched donors. In all cases, donor chimerism led to resolution of intestinal complications and infections, growth improvement, and correction of the immunodeficiency.

PATIENTS WITH IMMUNODEFICIENCY–CENTROMERIC INSTABILITY–FACIAL DYSMORPHISM (ICF) SYNDROME SHOW PERICENTROMERIC ANOMALIES OF CHROMOSOMES 1, 16, OR 9 IN MITOGEN-STIMULATED LYMPHOCYTES\(^{1} \) THAT CONSIST OF WHOLE-ARM DELETIONS, MULTIBRANCHED CHROMOSOMES, TRANSLOCATIONS, ISOCROMOSOMES, AND DECONDESMATION OF HETEROCROMATIC REGIONS ADJACENT TO THE CENTROMERES (JUXTACENTROMERIC HETEROCHROMATIN) OF THESE CHROMOSOMES AS A RESULT OF HYPMETHYLATION\(^{2} \) (Fig 1). ICF syndrome is the only immunodeficiency described to be caused by epigenetic dysregulation rather than single gene defects critical for lymphoid development or signaling. Some patients have mutations in a DNA methyltransferase gene, \( \text{DNMT3B} \),\(^{1} \) but others have no mutations identified.\(^{4} \) Clinical features include absent or low levels of serum immunoglobulins and facial anomalies.\(^{5,6} \) Primary ICF syndrome B cells have defective peripheral terminal B-cell differentiation, which contributes to agammaglobulinemia.\(^{7} \) Although hypogammaglobulinemia or agammaglobulinemia is most consistently described in patients with ICF syndrome, a number of these patients suffer significant infection associated with T-cell immunodeficiency, namely \textit{Pneumocystis jiroveci} pneumonia,\(^{8} \) viral pneumonia,\(^{9} \) persistent fungal infection, or viral enteritis, which lead to malabsorption and growth failure. Antiviral and antibacterial prophylaxis may be helpful, but in 1 study group, 18 (41%) of the 44 patients with ICF syndrome died by early adulthood, predominantly as a result of infection or related complications and lymphoid malignancy (C.W., unpublished observation).

Hematopoietic stem cell transplantation (HSCT) can cure severe combined immunodeficiencies and other

Abbreviations: ICF, immunodeficiency–centromeric instability–facial dysmorphism syndrome; HSCT, hematopoietic stem cell transplantation; \( \text{DNMT3B} \)
immunodeficiencies \textsuperscript{10,11} by replacing recipient with normal donor hematopoietic stem cells. Although patients with ICF syndrome suffer multisystem disease, death usually results from severe infection or sequelae such as bronchiectasis \textsuperscript{8}; therefore HSCT may be an attractive therapeutic option. Here we report on 3 unrelated patients who underwent HSCT to treat ICF syndrome.

**CASE REPORTS**

**CASE 1.** Patient 1 presented with recurrent chest infections after prolonged interstitial pneumonitis in early infancy, which resolved on antibiotic treatment including cotrimoxazole, although \textit{P. jiroveci} was not considered as a diagnosis. She was agammaglobulinemic (Table 1). Cytogenetic studies showed structural aberrations consistent with ICF syndrome; no mutation was identified in \textit{DNMT3B}, but hypomethylation of \textit{α} satellites was demonstrated, which correlates with another subset of patients with ICF syndrome. \textsuperscript{4} She received immunoglobulin replacement and antimicrobial prophylaxis. She persistently excreted small round structured virus from feces, accompanied by growth failure. At the age of 2.25 years, she underwent HSCT with marrow from a 10/10 HLA-matched unrelated donor after conditioning with alemtuzumab, fludarabine, and melphalan. Cyclosporine was given for graft-versus-host disease (GvHD) prophylaxis. On day 3 she developed severe vomiting and diarrhea with dehydration and acidosis, followed by capillary leak and increasing oxygen requirement. She required ventilation for 3 days and drainage of bilateral pleural effusions. Whole-blood genetic analysis on day 20 showed gain of donor and loss of recipient alleles. She cleared small round structured virus from feces. Cyclosporine was discontinued at 3 months, but the level of donor alleles, measured on whole blood, slipped to 15%. After an unconditioned boost infusion of marrow from the original donor, donor chimerism improved to 100% in all cell lineages and has remained at this level. Subsequently, \textit{Candida} nail and mouth infection resolved with antifungal treatment, and lobar pneumonia resolved with intravenous antibiotics, although no organism was grown. Immunoglobulin was discontinued at 2.5 years after HSCT; she has made normal antibody responses to vaccinations (Table 1). She has had persistent ulceration of the tongue; biopsy of the affected site has not demonstrated GvHD or fungal infection. Her growth has improved, and her weight has risen from 0.4th to the 50th percentile. There has been no evidence of autoimmunity after HSCT.

**CASE 2.** At 4 months of age, patient 2 had episodic diarrhea...
and vomiting that settled spontaneously, followed by an upper respiratory tract infection at 6 months. At 7 months of age, she developed pneumonia. She was neutropenic, anemic, and agammaglobulinemic. Cytogenetic analysis demonstrated classic chromosomal anomalies; heterozygous mutations in DNMT3B (Asp809Gly and Val605Ala) confirmed the diagnosis of ICF syndrome. She commenced immunoglobulin replacement and prophylactic cotrimoxazole. At 18 months of age, she underwent HSCT. Routine bronchoalveolar lavage revealed P jiroveci, which was treated with cotrimoxazole. She received marrow from a 10/10 HLA-matched unrelated donor after conditioning with ATG-Fresenius S (Fresenius, Bad Homburg, Germany), busulphan, and cyclophosphamide. Cyclosporine and methotrexate were given for GvHD prophylaxis. Whole-blood genetics on day 27 demonstrated 97% donor alleles. Cyclosporine was discontinued at 6 months and intravenous immunoglobulin at 9 months. At 18 months after HSCT, she made specific immunoglobulin G responses to vaccinations (Table 1).

She has experienced unstable thyroid function since the transplantation with positive antithyroid autoantibodies and biochemical evidence of hypothyroidism and hyperthyroidism. The donor’s thyroid status is unknown.

**CASE 3.** Patient 3 was born at term from consanguineous parents. A previously affected sister had died as a result of infection at 3 years of age. The diagnosis was confirmed postpartum by cytogenetic and DNMT3B mutational analysis, which demonstrated homozygous 2397-11G→A. The patient was agammaglobulinemic; immunoglobulin supplementation was started at 2 months. At 2 years he suffered Campylobacter lari infection that resulted in prolonged diarrhea, postinfectious enteropathy, and growth failure. Intestinal symptoms persisted in the absence of infections. At 4 years of age, HSCT was performed from an HLA-identical sister (the donor was homozygous for wild-type DNMT3B) after conditioning with busulphan and cyclophosphamide. GvHD prophylaxis consisted of cyclosporin and methotrexate. Full donor chimerism was demonstrated from day 30 onward. Cyclosporin was discontinued 5 months after HSCT. The diarrhea resolved, and he thrived; his height has risen from less than the 0.4th centile before HSCT to the 2nd to 9th centiles after HSCT, and his weight remains between the 2nd and 9th centiles. He made good responses to vaccination 4 months after HSCT after discontinuation of immunoglobulin replacement (Table 1); pneumococcal vaccination was not given at this point. Vitiligo developed 1 year after HSCT, and Streptococcus pneumoniae menigitis from which he recovered, albeit with the sequelae of serious hearing loss. Eighteen months after HSCT, the donor developed autoimmune hypothyroidism and, subsequently, so did the patient; both of them responded to thyroxine supplementation. At the time of this writing, it has been 2 years after HSCT, and the patient is well.

**DISCUSSION**

These cases demonstrate that HSCT corrects ICF syndrome-associated immunodeficiency. All the patients were hypogammaglobulinemic and had normal immunoglobulin levels after HSCT, with normal responses to vaccine antigen, including polysaccharide antigen. Furthermore, some B cells demonstrated the memory B-cell phenotype and evidence of class switching, which is not found with ICF syndrome.

Patients 1 and 3 experienced sustained infection-induced enteritis before HSCT. T-cell immune reconstitution led to resolution of gut symptoms and resumption of normal growth. Patients 2 and 3 experienced autoimmune phenomena, described after HSCT but not described with ICF syndrome. Patient 3 was euthyroid before HSCT; disease was probably donor derived. Nevertheless, vitiligo developed after HSCT. It is not possible to deduce with such a small number of patients whether ICF syndrome is genuinely a risk factor for the development of autoimmunity after HSCT similar to that seen in cases of mixed chimerism after HSCT for Wiskott-Aldrich syndrome.

Although ICF syndrome is not caused by an intrinsic hematopoietic stem cell defect, abnormalities in gene expression critical for B-cell immunoglobulin isotype switching and lymphocyte activation and migration have been observed. Agammaglobulinemia is the most obvious immunologic abnormality at presentation, but a number of patients experience complications that are more suggestive of significant T-cell immunodeficiency. Recent studies in murine models seem to support the impact of ICF mutations on T-cell development. Critically, donor B-cell reconstitution will be required in ICF syndrome to achieve full immune reconstitution, with B-cell function.

**CONCLUSIONS**

These 3 patients tolerated either full myeloablative conditioning with busulphan/cyclophosphamide or reduced-intensity conditioning with fludarabine/melphalan, with no significant demonstrable toxicity. HSCT can cure humoral and cellular ICF syndrome-associated immunodeficiency. Given the poor survival without HSCT and these results, HSCT should be considered for patients with ICF syndrome with clinical and/or laboratory evidence of T-lymphocyte dysfunction. Additional studies may ascertain the risk of developing autoimmune-mediated sequelae after HSCT, emphasizing the importance of international databases that record the outcome of patients with these rare syndromes.
REFERENCES

Severe Varicella Caused by Varicella-Vaccine Strain in a Child With Significant T-Cell Dysfunction

Patrick Jean-Philippe, MDa,b, Abigail Freedman, MDa,b, Mary Wu Chang, MDa,c, Sharon P. Steinbergd, Anne A. Gershond, Philip S. LaRussa, MDb, William Borkowsky, MDa,b

aDepartment of Pediatrics, bSaul Krugman Division of Infectious Diseases and Immunology, and dRonald O. Perelman Department of Dermatology, New York University School of Medicine, New York, New York; dDivision of Infectious Diseases, Department of Pediatrics, Columbia University Medical Center, New York, New York

Financial Disclosure: Dr Gershon consults and lectures for Merck and GlaxoSmithKline on varicella vaccine. The other authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

In March 1995, the US Food and Drug Administration approved a live attenuated varicella vaccine for use in healthy children 12 months to 12 years old. We report here an 18-month-old girl with cell-mediated immunodeficiency who developed a severe vaccine-associated rash and clinical evidence of vaccine-associated pneumonia 1 month after inadvertent receipt of varicella vaccine.

IN MARCH 1995, after >20 years of testing, the US Food and Drug Administration approved a live attenuated varicella vaccine (LAVV) for prevention of varicella in susceptible healthy individuals 12 months of age or older.1 This vaccine, the Oka strain, was developed in Japan by Takahashi et al2 and had been used there since 1974.3 Because of the severe morbidity of varicella in immunocompromised persons, early US trials were conducted in leukemic patients as well as in healthy children, with excellent efficacy and safety profiles resulting.4,5 However, vaccine use in many immunocompromised populations has remained contraindicated or restricted because of fear that vaccine virus might cause disease in these patients.6–9 Here we report 1 such instance: a case of severe disseminated vaccine-strain varicella in a child with an uncharacterized cell-mediated immunodeficiency.

CASE REPORT

An 18-month-old black girl was admitted to Bellevue Hospital (New York, NY) with a 4-day history of fever and increasing numbers of papulovesicular/pustular skin lesions. The rash was first noted on the patient’s trunk 6 days before admission when she began having intermittent fevers. At that time, she was seen by her primary physician as an outpatient and evaluated by a pediatric dermatologist. Since that time, the lesions had become generalized, and on admission they covered her entire body including her scalp, palms, and soles. A culture from blood drawn 4 days before admission did not grow any bacteria, but a culture from 1 of the pustules revealed infection with methicillin-resistant *Staphylococcus aureus*. Three days before admission, she was started on benzathine penicillin and erythromycin, but the rash continued to progress. On the evening before admission she was brought to the emergency department because of a fever of 103°F.

Medical History

This patient’s medical history was significant for preterm delivery at 32 weeks’ gestation via cesarean section, secondary to maternal prolactinoma and polyhydramnios. At birth she was noted to have choanal atresia and a patent ductus arteriosus; a choroid plexus cyst was also noted on computed tomography of the head, and bilateral renal calculi were seen on ultrasound. The result of
prenatal testing of the patient’s mother for HIV antibody was negative, and the result of testing of the patient for HIV antibody as part of the state’s newborn screening program was also negative. In the early postnatal period she required both a tracheostomy and gastrostomy-tube placement. Although her anomalies suggested a congenital syndrome, a specific diagnosis was not established at birth or through subsequent investigation.

The finding of persistent lymphopenia on serial lymphocyte count and the absence of a thymic shadow on chest radiograph raised the suspicion of DiGeorge anomaly, but results of a fluorescent in situ hybridization study to detect the 22q11.2 deletion were negative. However, initial determination of lymphocyte subsets revealed severe depletion of both CD4 and, particularly, CD8 subsets, with normal or increased B cells (CD19) and natural killer cells (CD56). Lymphocyte-proliferation assays also revealed absent proliferation to standard mitogens. This trend would persist, albeit with some normalization of CD4 percentage, until this admission (Table 1). In addition to her deficits in cellular immunity, the patient also had severe humoral dysregulation. Although her immunoglobulin (Ig)G, IgM, and IgA concentrations were within normal limits for her age, she failed to produce protective antibody titers to tetanus toxoid and Haemophilus influenzae type B after immunization. Further workup of her immunodeficiency included testing for genotypes of severe combined immunodeficiency, such as ZAP-70 deficiency and associated promoter mutations, all of which were unremarkable. Because of the nursing care her tracheostomy required, she was placed in a chronic care facility.

Before this admission, the patient had been hospitalized multiple times. During several admissions for exacerbations of chronic reactive airway disease, test results of nasopharyngeal aspirates were positive for respiratory syncytial virus. On 1 of these admissions, at 9 months of age, the patient suffered an exacerbation of her chronic ichthyosiform rash that was manifested by diffuse erythroderma, alopecia, and electrolyte disturbances. Quantitative Igs at that time revealed markedly elevated serum IgE (36 000 IU); however, her clinical course was not felt to be compatible with Job syndrome because of absent history of deep organ abscess. A diagnosis of Netherton syndrome was considered but not supported by the results of skin biopsy and hair-shaft examination. Furthermore, molecular testing results were negative for mutation on SPINK5. The diagnosis of Omenn syndrome was also considered, but testing for RAG1 and RAG2 was not performed. Despite the gastrostomy tube, she continued to suffer from failure to thrive and weighed only 8.6 kg at 18 months of age, with profound developmental delay. No cause for her immunodeficiency was ever elucidated, but because of this condition, her physician advised that she receive no live vaccines. Nevertheless, 5 weeks before her admission, she was given Varivax (Merck and Co, West Point, PA), the LAVV.

### Hospital Course

On admission, the patient was awake, alert, responsive, and moving all extremities. She was afebrile but hypertensive, tachycardic, and tachypneic and had a blood pressure of 120/80 mm Hg, a heart rate of 150 beats per minute, and respiratory rate of 45/minute on 28% O2 via a tracheostomy collar. There were generalized erythematous-based vesicles and pustules, including on her palms, soles, scalp, and trunk, with the highest concentration being in the genital area. Her breath sounds were coarse, with occasional wheezes; her liver was palpable 3 cm below the costal margin, but laboratory testing on admission revealed a normal complete blood count, chemistry, and liver function.

The patient was admitted with a preliminary diagnosis of staphylococcal pustulosis, but because of the mixed nature of the lesions and the history of varicella immunization, treatment was begun with intravenous vancomycin, intravenous ceftriaxone, and intravenous acyclovir. A 3-mm punch biopsy was performed by the pediatric dermatology service, 2 vesicles on the dorsum of the left foot were unroofed for bacterial, fungal, and

### Table 1: Longitudinal Immunologic Data

<table>
<thead>
<tr>
<th>Age, mo</th>
<th>Flow Cytometry, %</th>
<th>Lymphoproliferative Assay, cpm × 10⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CD4⁺</td>
<td>CD8⁺</td>
</tr>
<tr>
<td>2.9</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>3.1</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>8.2</td>
<td>72</td>
<td>4</td>
</tr>
<tr>
<td>9.4</td>
<td>70</td>
<td>15</td>
</tr>
<tr>
<td>9.4a</td>
<td>54</td>
<td>16</td>
</tr>
<tr>
<td>16.3</td>
<td>30</td>
<td>12</td>
</tr>
</tbody>
</table>

*Shown are longitudinal flow-cytometric and lymphoproliferative-assay data from the patient featured in the report. Despite improvement in the number and distribution of the T-cell subsets, polyclonal proliferation in response to antigens remained markedly impaired. — indicates that tests were not performed.

a Phytohemagglutinin (reference value: >70 000 cpm).
b Pokeweed mitogen (reference value: >25 000 cpm).
c Tetanus toxoid (reference value: >3-fold increase over homologous control wells).
da Lymph node biopsy specimen.
viral cultures, and a Tzanck smear was obtained; the results of these tests were negative. On hospital day 4, fluid obtained from 1 of the vesicles was sent to Columbia University for a diagnostic polymerase chain reaction (PCR) and viral strain determination.

New lesions continued to appear for >1 week, for a total of >14 days after initial appearance of the rash, and the patient continued to have low-grade fevers and mild respiratory distress. A chest radiograph on hospital day 4 revealed bilateral patchy consolidation. Two punched-out ulcers appeared on her left flank and another on her left frontal scalp; all 3 were >1 cm in diameter and surrounded by 3 to 4 cm of erythema and induration. On hospital day 4, the pathology report on the biopsy reported results consistent with a herpessimplex virus infection, although differentiation between herpes simplex virus and varicella-zoster virus (VZV) was not possible. On the 4th hospital day, the PCR was reported positive for vaccine-strain (Oka) varicella. After the 9th hospital day, new lesions gradually diminished, and by the 17th hospital day, the punched-out ulcers began to fill with granulation tissue. Although the patient’s hospital course was complicated by *S aureus* bacteremia and candidemia, as well as methicillin-resistant *S aureus* superinfection of the primary lesions, her varicella continued to improve after this point. After a lengthy and difficult hospital course, she was discharged back to her chronic care facility on the 55th hospital day.

The patient was subsequently readmitted on several occasions for ulcerated skin lesions and respiratory illness. VZV was never isolated from any specimen. Her lymphocyte-proliferation responses remained severely depressed, and she eventually died at 2 years, 2 months of age from respiratory failure.

**DISCUSSION**

In March 1995, the US Food and Drug Administration approved an LAVV (Varivax) after almost 2 decades of clinical trials had shown it to be safe, effective, and immunogenic. The Advisory Committee on Immunization Practices initially recommended that the use of this vaccine be restricted to healthy children 12 months to 12 years of age.1 Subsequent updated recommendations of the Advisory Committee on Immunization Practices allowed its use to be considered in asymptomatic or mildly symptomatic HIV-infected children in Centers for Disease Control and Prevention class N1 or A1 with an age-specific CD4+ T-lymphocyte percentage of ≥25%,11 and it was recently shown to be safe and immunogenic in HIV-infected children with a CD4 percentage between 15% and 24%.12

Although the genetic basis of vaccine-strain attenuation is still being elucidated, specific phenotypic correlates have been proposed. In experiments using the human severe combined immunodeficiency mouse model, the vaccine strain has shown decreased ability to replicate in skin compared with the wild-type (WT) strain.13 It has also demonstrated decreased transmissibility compared with WT; secondary transmission of varicella-vaccine strain from a vaccine recipient with rash to close contacts has been documented in only a few instances14–17 and has been significantly less than what would be expected from WT strain. Finally, evidence indicates decreased propensity to reactivate as herpes zoster,18,19 although such cases have been reported.17,20

The mechanism by which the vaccine strain confers protection has yet to be elucidated conclusively; however, it has been speculated that production of nonreplicating immunogenic viral particles at the site of inoculation, as well as decreased infectivity for the skin compared with the WT strain, allows time for the induction of an adaptive VZV-specific cell-mediated immune response before viremia occurs.21 Previous studies have shown that the extent of VZV-specific T-cell response 3 days into the course of the disease determines the extent of the rash in WT infection.22,23 Patients with impaired cellular immune responses can have viremia after vaccine administration. However, in most clinical studies performed before and after licensure, varicella vaccine has proved to be safe and effective in healthy recipients.4,17,24,25 A recent study performed in 3 active surveillance sites showed significant reduction in the number of new cases of varicella after introduction of the vaccine.25 Subsequent studies have also demonstrated a substantial health care impact of the vaccine, including significant reduction in VZV-related hospitalizations,26,27 emergency department visits,26 medical expenditures,27 and mortality in all age groups.28 The latter effect was highest in the 1- to 4-year-old age group, with 92% reduction in varicella-induced mortality.

LAVV has also shown a favorable profile in some categories of immunocompromised patients, especially leukemic children for whom most of the data are available. As in healthy children, the most common adverse event after immunization is a mild vaccine-associated rash, usually maculopapular and vesicular, that occurs ~1 month after immunization in 5% of vaccine recipients who are not receiving chemotherapy and in up to 50% of leukemic children who are receiving chemotherapy at the time of immunization.29,30 It also protected 86% of leukemic vaccine recipients from household exposure to chickenpox.31,32 Using Varilrix (SmithKline Beecham, Rixensart, Belgium), Leung et al33 reported a good safety and immunogenicity profile on pediatric patients with hematologic cancers in maintenance phase or solid tumors 3 to 6 months after discontinuation of chemotherapy.34 Recent data on a limited number of pediatric solid organ (kidney, liver, and intestine) transplant recipients suggested that varicella vaccine is also safe and immunogenic in this immunosuppressed population.34,35

The patient described in this report had profound
impairment of cellular immune function as well as impaired humoral responses. Intensive genetic and immunologic workup never identified an etiology for her immunodeficiency. At 17 months of age the patient inadvertently received LAVV and, 28 days later, developed a vesicular eruption that consisted of hundreds of new lesions that appeared as late as 2 weeks after onset of the eruption despite antiviral therapy. The appearance and evolution of these lesions, described above, were consistent with the appearance of WT VZV in immuno compromised patients; although the patient had lesions of variable sizes, the eruption was monomorphic (ie, lesions evolved synchronously), in contrast to varicella in a normal individual on whom the lesions evolve asynchronously (new lesions mixed with healing, older lesions). In addition, the concurrent tachypnea, oxygen requirement, and appearance on a chest radiograph of bilateral patchy infiltrates are suggestive of varicella pneumonia. Previous reports have used similar criteria for diagnosis of varicella pneumonia in immunosuppressed patients. A VZV PCR using restriction fragment length polymorphism revealed a pattern consistent with Oka vaccine virus strain: presence of a novel Bgl II site in gene 53 and absence of a Pst I site in gene 38. Review of the literature reveals few reports of serious adverse effects associated with varicella-vaccine strain in immunocompromised patients: a 19-year-old adolescent with primary sclerosing cholangitis and lymphopenia who developed a severe varicella-vaccine-induced vesicular eruption after receipt of varicella vaccine; a 13-month-old boy subsequently diagnosed with adenosine deaminase deficiency who developed hepatitis and a generalized vesicular eruption caused by the Oka vaccine strain; a 16-month-old boy with HIV infection and a CD4 count of 8 cells per mm who developed Oka vaccine-strain–induced pneumonitis; a 5-year-old boy with cerebral palsy and reactive airway disease who received a dose of varicella vaccine 7 days after completing a steroid taper and restarted on steroid therapy 8 days after receiving the vaccine who developed a rash and pneumonia, with the Oka vaccine strain recovered from his endotracheal secretions; an 11-year-old girl diagnosed with a natural killer T cell deficiency who developed rash and severe pneumonitis caused by Oka strain infection 5 weeks after receipt of LAVV, and a 1-year-old boy vaccinated with the Oka strain shortly before diagnosis with a neuroblastoma that required intensive chemotherapy. Although the 1-year-old boy did not contract acute varicella, he subsequently developed chronic disseminated herpes zoster that was shown to be caused by the vaccine strain. A 24-year-old resident of a developmental center with panhypopituitarism who was receiving physiologic doses of daily steroids developed a diffuse febrile vesicular rash and roentgenographic evidence of right basilar pneumonia 18 days after receipt of varicella vaccine, although the presence of vaccine strain was not demonstrated.

CONCLUSIONS
This case report illustrates the fact that, although VZV vaccine has been proven to be safe in immunocompetent patients, it is potentially dangerous in patients with altered immunity, especially those with severely suppressed cell-mediated immunity, in whom it can produce long-lasting severe generalized eruption or organ dissemination. Although it may not be cost-effective to perform routine immunodeficiency screening in all apparently healthy children who present for vaccination with LAVV, particular attention should be paid to clues such as abnormal anthropomorphic data to detect patients who might be at increased risk of immunodeficiency and, therefore, at increased risk of adverse effects from vaccination with LAVV. This case also highlights the fact that if immunodeficiency is suspected, assessment of function in addition to phenotype should be conducted before excluding it.

REFERENCES


16. Brunell PA, Argaw T. Chickenpox attributable to a vaccine virus contracted from a vaccinee with zoster. *Pediatrics.* 2000;106(2). Available at: www.pediatrics.org/cgi/content/full/106/2/e28


Neonatal Hyperparathyroidism and Pamidronate Therapy in an Extremely Premature Infant

Lisa Fox, MB, ChB, FRACP; Joel Sadowsky, MB, ChB, FRACP; Kevin P. Pringle, MB, ChB, FRACS; Alexa Kidd, MBBS, MRCP, MSc; Jean Murdoch, MD, FRCP(C); David E. C. Cole, MD, PhD, FRCP(C); Esko Wiltshire, MD, FRACP; Departments of Paediatrics and Child Health and Radiology and Central Regional Genetics Service, Wellington Hospital, Capital and Coast District Health Board, Wellington, New Zealand; Departments of Surgery, Obstetrics and Gynaecology, and Paediatrics and Child Health, Wellington School of Medicine and Health Sciences, University of Otago, Wellington, New Zealand; Departments of Laboratory Medicine and Pathobiology, Medicine, and Pediatrics, University of Toronto, Toronto, Ontario, Canada

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

We describe the use of pamidronate to control marked hypercalcemia in an extremely premature infant with neonatal hyperparathyroidism that resulted from an inactivating mutation (R220W) of the calcium-sensing receptor. Despite improvement in bone mineralization and subsequent parathyroidectomy with normalization of the serum calcium level, the combination of chronic lung disease, osteomalacia, and poor thoracic cage growth ultimately proved fatal. Pamidronate therapy seems to be safe in the short-term and effective in helping control hypercalcemia even in the very premature infant, allowing for planned surgical intervention when it becomes feasible.

Neonatal hyperparathyroidism (NHPT) is a rare disorder that can be caused by loss-of-function mutations in the calcium-sensing receptor (CaSR). Most reported cases have been of term infants; there has been 1 report of an affected premature infant but no reports of extremely premature infants. Total parathyroidectomy may be life saving for severely affected infants, although novel medical therapies are being investigated. A recent report has demonstrated the effectiveness of bisphosphonate (pamidronate) therapy in reversing severe hypercalcemia in term infants with recessive neonatal severe hyperparathyroidism (NSHPT), which allows parathyroidectomy to be delayed until the infant is clinically stable.

The CaSR establishes the “set point” for calcium homeostasis. Loss-of-function mutations lead to a higher set point for serum calcium level and cause a spectrum of disease including (1) the most severe form (NSHPT), with life-threatening hypercalcemia and bone disease secondary to demineralization; (2) NHPT with moderate hyperparathyroidism and hyperparathyroid bone disease, and (3) familial hypocalciuric hypercalcemia (FHH), which is a relatively benign autosomal-dominant disorder with asymptomatic hypercalcemia.

Bisphosphonates have been used in neonates with hypercalcemia caused by subcutaneous fat necrosis and with osteogenesis imperfecta, but we could find no reports of use of these agents in premature infants. Here we report an extremely premature infant with NHPT and severe bone disease, outline the difficulties in management, and describe the safe short-term use of pamidronate at early gestation.

PATIENT REPORT AND METHODS

The patient, a girl, is the first child of healthy, nonconsanguineous parents and was delivered by emergency cesarean section at a gestational age of 27 weeks 4 days after an acute antepartum hemorrhage. One dose of betamethasone was administered before delivery. Her birth weight was 1230 g (90th centile), head circumference was 27.6 cm (97th centile), and birth length was 37 cm (90th centile). She was intubated at delivery and given surfactant. The initial chest radiograph (Fig 1) showed skeletal changes consistent with hyperparathyroidism and pulmonary changes consistent with respiratory distress.

Key Words: neonate, bisphosphonate, hypercalcemia, calcium-sensing receptor, familial hypocalciuric hypercalcemia, neonatal hyperparathyroidism

Abbreviations: NHPT, neonatal hyperparathyroidism; CaSR, calcium-sensing receptor; NSHPT, neonatal severe hyperparathyroidism; FHH, familial hypocalciuric hypercalcemia; PTH, parathyroid hormone; CGA, corrected gestational age

www.pediatrics.org/cgi/doi/10.1542/peds.2006-3209
doi:10.1542/peds.2006-3209
Accepted for publication May 4, 2007
Address correspondence to Lisa Fox, MB, ChB, Neonatal Services, Royal Women’s Hospital, 132 Grattan St, Carlton, Victoria 3053, Australia. E-mail: lisa.fox@rwh.org.au
PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
A family history of FHH included the paternal grandmother, and several members of the paternal family underwent parathyroidectomy as adults. The patient’s mother had a serum total calcium level (corrected for albumin) of 2.5 mmol/L (adult reference range: 2.25–2.50 mmol/L), serum phosphate level of 1.38 mmol/L (reference range: 0.95–1.60 mmol/L), and parathyroid hormone (PTH) level of 1.3 pmol/L (reference: 0–6 pmol/L). However, the patient’s father had a corrected calcium level of 3.06 mmol/L, serum phosphate level of 0.87 mmol/L, and PTH level of 2.7 pmol/L, a fact that had not been recognized previously.

Hypercalcemia was first noted at 15 hours of age, with total a corrected serum calcium level (corrected for albumin) of 2.5 mmol/L (adult reference range: 2.25–2.50 mmol/L), serum phosphate level of 1.38 mmol/L (reference range: 0.95–1.60 mmol/L), and parathyroid hormone (PTH) level of 1.3 pmol/L (reference: <6 pmol/L). However, the patient’s father had a corrected calcium level of 3.06 mmol/L, serum phosphate level of 0.87 mmol/L, and PTH level of 2.7 pmol/L, a fact that had not been recognized previously.

Hypercalcemia was first noted at 15 hours of age, with total a corrected serum calcium level of 3.08 mmol/L (neonatal reference range: 2.25–2.70 mmol/L), phosphate level of 2.05 mmol/L (reference range: 1.55–2.65 mmol/L), alkaline phosphatase level of 203 U/L (reference range: 0–400 U/L), PTH level of 233 pmol/L (reference: <6 pmol/L), urine calcium/creatinine ratio of 1.33 (reference: <1.96 mmol/mmol in infancy), and 25-hydroxyvitamin D level of 15.3 µg/L (18–56 µg/L, day 2).

**Treatment and Progress**

The patient’s hypercalcemia was managed initially with hyperhydration, frusemide, and steroids, with minimal effect. She also received treatment with 400 IU of vitamin D daily. On day 7, treatment with intravenous pamidronate (20 mg/m²) was initiated. Six doses of 20 to 30 mg/m² were given over the next 6 weeks with the aim of performing parathyroidectomy once it was technically feasible and she was clinically stable. Pamidronate was given in 5% (wt/vol) dextrose over 4 hours with a maximum concentration of 240 µg/mL. Her serum calcium level reduced significantly, by 7% to 28%, with each dose (Fig 2).

The radiographic appearances initially worsened, with a visible increase in subperiosteal bone resorption and intracortical tunneling. A skull radiograph taken on day-of-life 10 is shown in Fig 3. Improvement in radiographic bone density was noted 2 weeks after the first pamidronate infusion, with evidence of healing of femoral metaphyseal fractures at 1 month (2 infusions given) and subperiosteal resorption of bone subsequently resolved. There was significant periosteal reaction along the long tubular bones, with increased density and thickness, from 7 weeks of age (when 5 infusions had been given), which is earlier than usually seen for physiologic periosteal reaction of the newborn (as demonstrated in Fig 4). Radiographs were reviewed by 2 radiologists, each with >10 years’ experience in neonatal radiograph interpretation at a tertiary-level neonatal unit.

**FIGURE 1**

Chest radiograph on day 1 demonstrating short ribs with anterior rib fractures, significant osteopenia outside the normal appearance for gestation, and subperiosteal resorption of bone, intracortical tunneling, and poor definition of trabeculae (as classically observed in hyperparathyroidism). Lung fields are small. The diffuse, slightly coarse infiltrate present throughout with areas of punctate luencies suggests early pulmonary interstitial emphysema on a background of pulmonary hypoplasia and mild respiratory distress syndrome.

**FIGURE 2**

Effect of treatment on calcium and PTH. A, Change in PTH over time (the reference interval [<6 pmol/L] is shaded). B, Change in serum calcium (the reference interval is shown with lighter shading). Pamidronate infusions (20 or 30 mg/m²) are indicated by gray bars, and the parathyroidectomies (days 55 and 105) are indicated by arrows.
We did not observe any immediate adverse effects from the pamidronate; specifically, there was no hypomagnesemia, vomiting, or lymphocytopenia, no evidence of a febrile reaction with the first dose, and no respiratory distress associated with infusion, as is seen occasionally in infants with osteogenesis imperfecta.

Parathyroidectomy was performed at 8 weeks of age (corrected gestational age [CGA]: 35 weeks). Globular collections of brown fat appeared very similar to neonatal parathyroid tissue. Four (of 9) pieces of excised tissue were confirmed histologically to be hyperplastic parathyroid tissue. Her calcium and PTH levels decreased initially but rose to previous levels within 48 hours. Repeat surgical exploration at 42 weeks' CGA removed a fifth parathyroid gland, which resulted in rapid and permanent reduction in her calcium and PTH levels (Fig 2). Supplemental calcium (up to 1.7 mmol every 6 hours and calcitriol (0.25 μg daily) were given to prevent symptomatic hypocalcemia.

Despite pamidronate treatment, growth of her thorax was poor. She developed progressive thoracic deformity, limb fractures, and severe chronic lung disease. She was unable to sustain adequate respiratory function without ventilator support for >24 hours. At 5 months of age, bronchoscopy demonstrated tracheomalacia, and a thoracic computed tomography scan showed areas of hyperinflation and atelectasis consistent with chronic lung disease. Given her poor respiratory prognosis, intensive care was withdrawn at 5.5 months of age (CGA: 9 weeks postterm) and she died as a result of respiratory failure.

Molecular Analysis
Genomic DNA was extracted from leukocytes, the entire coding sequence of the \( \text{CaSR} \) gene (including the intron-exon splice junctions) was amplified by polymerase chain reaction, and the amplicons were subjected to denaturing high-performance liquid chromatography (dHPLC) to identify heterozygosity. To ensure that there were no mutations with homoallelic homozygosity, each fragment was mixed with an equal amount of amplified control DNA, heated to 95°C, and allowed to reanneal and form heteroduplexes. dHPLC analysis identified heterozygosity in exon 4, and bidirectional sequencing confirmed a single C-to-T transition at nucleotide position 658, which predicts heterozygosity for an R220W missense mutation. Search of the Calcium Sensing Receptor Database (available at www.casrdb.mcgill.ca) revealed that this mutation had been reported previously in 3 other families.

DISCUSSION
To our knowledge, this is the first report of NHPT and use of pamidronate in an extremely preterm infant. In its most severe form, NSHPT, this disease presents with failure to thrive, lethargy, poor feeding, ileus, and hypotonia. Rapid progression of skeletal demineralization leads to multiple fractures and thoracic deformities. The mortality rate has been reported at up to 50% in some series, because of severe hypercalcemia, respiratory failure secondary to thoracic deformation, and/or respiratory infection. In survivors, neuropsychological development may also be significantly affected.

In NHPT in a term infant, the hypercalcemia is less severe, as is the bone disease. However, our extremely premature patient had a skeletal phenotype that more closely resembled that seen in NSHPT despite control of hypercalcemia with pamidronate infusions. Whether this is related to the specific genotype or also reflects interplay with external exacerbating factors deserves comment. Inactivating \( \text{CaSR} \) mutations cause a wide
spectrum of disease that ranges from stillbirth and life-threatening neonatal disease to neonatal skeletal demineralization that resolves with time and even to entirely asymptomatic FHH. Within our patient’s family, there was no history of NHPT, but several adults had undergone parathyroidectomy for symptomatic hypercalcemia and elevated serum PTH concentrations, suggesting a more severe phenotype than seen in many other families with FHH. Molecular analysis of CaSR revealed a heterozygous R220W mutation, previously associated with FHH and NHPT. The substitution of the arginine at position 220 by tryptophan occurs in the portion of the amino-terminal extracellular domain of the CASR protein that putatively participates in calcium binding. In attempting to explain the occurrence of NHPT in their patient, Schwarz et al suggested that R220W interrupts ligand binding and exerts a dominant-negative effect, as has been the case in other mutations. Using a transiently infected human embryonic kidney (HEK) cell-culture system to explore the functional properties of this mutation, D’Souza-Li et al found that the R220W mutant has a right-shifted calcium-response curve and a calcium set-point concentration that is more than threefold higher than that in controls (15.4 ± 0.5 vs 4.0 ± 0.1 mmol/L) but is expressed normally at the cell surface. Additional functional studies are needed. The CaSR is expressed in the placenta. In mice, inactivating mutations in the CaSR reduce transfer of calcium to the developing fetus. A similar effect in humans would add to skeletal demineralization (by further increasing fetal PTH release).

Whatever the molecular consequences of the R220W mutation, they cannot readily explain the clinical variability within this family. Transmission of FHH through the father is likely to be an important factor, as seen in other families. With paternal transmission, the maternal environment is one of normal calcium homeostasis. In utero, the fetal parathyroid gland and its partially inactivated CaSR senses the “normal” calcium level as low and stimulates fetal PTH release, thereby supporting increased fetal calcium at the expense of skeletal mineralization, which results in prenatal fractures and interferes with proper thoracic cage growth. A second factor may be mild vitamin D deficiency. Our patient had a low 25-hydroxyvitamin D level on day-of-life 2, consistent with longer-term maternal vitamin D insufficiency, which would be an additional stimulus on the parathyroid gland. Extreme prematurity in itself also may have affected this infant’s skeletal mineralization. Bony mineralization is slower for premature infants, who lose the opportunity to acquire the large percentage of minerals usually transferred in the third trimester. It has also been postulated that the altered hormonal and biochemical environment of preterm postnatal existence contributes to decreased bone formation and increased reabsorption.

Severe chronic lung disease (caused by prematurity, thoracic deformity, and ventilator dependency) and extra parathyroid tissue not identified during the initial surgery complicated our patient’s course. Seven percent of the normal population has >4 parathyroid glands. In this case, our patient’s hyperparathyroid state was prolonged for an additional 6 weeks until surgery could be repeated safely with a reasonable chance of success.

Total parathyroidectomy is the standard treatment for NSHPT and leads to significantly greater survival than subtotal parathyroidectomy or medical treatment. However, infants with milder NHPT and a heterozygous CaSR mutation may improve spontaneously, particularly if the only factor is an affected father but unaffected mother, when removal from the relatively hypocalcemic fetal environment allows the infant to obtain sufficient calcium to improve osteopenia.

Pamidronate has been used in term neonates in the management of NSHPT to treat severe hypercalcemia until parathyroidectomy could be performed. In our case, pamidronate was introduced early with an aim to not only treat hypercalcemia but also increase bony density and improve chest wall mechanics and thereby limit development of severe chronic lung disease pending parathyroidectomy. In this case, to our knowledge the first extreme preterm infant treated with pamidronate, it was effective in reducing severe hypercalcemia and led to improvement in bone density on radiographs. We did not observe adverse effects in doses of 20 to 30 mg/m² at intervals of 4 to 14 days in the short-term; unfortunately, it did not prevent development of thoracic deformity and ultimately fatal severe chronic lung disease.

CONCLUSIONS

NHPT remains a challenging disorder, particularly because the disease may be well advanced by delivery as a result of the unique interrelationship of the maternal-fetal dyad. Pamidronate showed some therapeutic benefit and did not have short-term adverse effects in this preterm infant, although use of this treatment could not overcome the impact of prematurity and development of chronic lung disease.

REFERENCES

4. Cole DEC, Janicic N, Salisbury SR, Hendy GN. Neonatal severe hyperparathyroidism, secondary hyperparathyroidism, and fa-
Atypical Tetanus in a Completely Immunized 14-Year-Old Boy

Kai König, MD, Hannelore Ringe, MD, Brigitte G. Dorner, PhD, Alexander Diers, MD, Birgit Uhlenberg, MD, Dominik Müller, MD, Verena Varnholt, MD, Gerhard Gaedicke, MD

*Children’s Hospital, Charité Universitätsmedizin Berlin, Campus Virchow-Klinikum, Berlin, Germany; Robert Koch Institute, Microbial Toxins, Center for Biological Safety, Berlin, Germany

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

We report the uncommon clinical course of tetanus in a completely immunized 14-year-old boy. His initial symptoms, which included a flaccid paralysis, supported a diagnosis of botulism. Preliminary mouse-test results with combined botulinum antitoxins A, B, and E, obtained from tetanus-immunized horses, backed this diagnosis. The change in his clinical course from paralysis to rigor and the negative, more specific, botulinum mouse test with isolated botulinum antitoxins A, B, and E, obtained from nonvaccinated rabbits, disproved the diagnosis of botulism. Tetanus was suspected despite complete vaccination. The final results of a positive mouse test performed with isolated tetanus antitoxin confirmed the diagnosis. Adequate treatment was begun, and the boy recovered completely.

TODAY, TETANUS IS a rare disease in countries with primary immunization programs. The reported incidence for adults and children is low, with an average of 43 cases per year in the United States,1 12 to 15 per year in the United Kingdom,2 and <15 per year in Germany.3 However, tetanus occurs occasionally despite complete vaccination status. In these cases, the clinical picture can be altered, which hampers accurate and timely diagnosis. Here we report the unusual clinical presentation of tetanus in a completely immunized 14-year-old German boy.

CASE REPORT

The patient was admitted to our hospital with a 1-day history of headache, left-sided ptosis, generalized paresthesia, and impaired vision. His oral mucous membranes were extremely dry. Three days before he had suffered from mild diarrhea, and the day before admission he had eaten grilled chicken with barbeque sauce. The parents recalled a tick bite 1 year ago and an accidental abrasion at the patient’s left knee 1 week before, when the boy scratched himself on the rough surface of wooden floorboards. He did not clean the wound, and at admission it was small (2–3 mm in diameter), dry, and in the process of healing. In his spare time, the patient, who is right-handed, used to work with lacquers and glue. His medical history was uneventful. He had had chicken pox at the age of 3 years and common upper respiratory tract infections in his infancy. His immunization schedule was up-to-date with initially 3 vaccinations in the first year of life and a tetanus booster 1 year before presentation. Communication with the patient’s doctor and with the manufacturer of the vaccine revealed that the booster was within the vaccine’s expiration date and that there were no reports of reduced quality of that production lot.

After admission, an antibiotic and antiviral treatment with ceftriaxone (200 mg/kg per day), clarithromycin (15 mg/kg per day), and acyclovir (30 mg/kg per day) was initiated for suspected early meningoencephalitis despite normal cerebrospinal fluid test results. The results of microbiologic and virologic examinations (cultures and polymerase chain reaction for bacteria and fungi, ameba, toxoplasmosis, Mycoplasma, Borrelia, Chlamydia, rabies, herpes simplex 1–2, HIV 1/2, cytomegalovirus, Epstein-Barr virus, enterovirus, early-summer meningoencephalitis, measles, varicella, influenza and parainfluenza 1–3, and parvovirus B19) in blood and cerebrospinal fluid were negative. Results of drug...
screening, an edrophonium-provocation test, and testing of autoimmune antibodies (antineutrophil cytoplasmic antibody, antinuclear antibody, antimitochondrial antibody, and antibodies against neurons, myelin, glycoprotein, and ganglioside) were negative.

On day 2, the patient’s condition deteriorated severely, with alternating hypopnea and tachypnea, anxiety, hyporeflexia, bilateral ptosis, oculomotor nerve palsy, photophobia, dysarthria, dysphagia, and flaccid paralysis of the trunk and lower limbs. The patient was transferred to the PICU. Repeated electroencephalography, neurophysiologic examinations, and cerebral MRI and magnetic resonance angiography were normal. Treatment with intravenous immunoglobulin (Gammaglobulin [Chiron-Behring GmbH & Co KG, Marburg, Germany], 2 g/kg = 100 g over 5 days intravenously) was started for suspected Guillain-Barré syndrome. Because an atypical botulism infection could not be completely excluded, equine botulinum antitoxin (Botulinus-Antitoxin [Chiron-Behring GmbH & Co KG, Marburg, Germany], 1 mL = 750 IU of botulinum antitoxin A, 500 IU of antitoxin B, and 50 IU of antitoxin E) was added on day 3. The antidote treatment was ceased at 350 of 500 mL because of an anaphylactic reaction. On the same evening he developed urine bladder dysfunction, carpopedal spasms, intermittent rigors of the upper limbs, and increasing rigors of the lower limbs and hypopnea. He received midazolam (0.04 mg/kg per hour intravenously), tetrazepam (1 mg/kg per day orally), and metamizole (80 mg/kg per day intravenously) to control spasms and pain. Hypopnea and apnea were treated with theophylline (initially 5 mg/kg, then 4 × 2.5 mg/kg per day intravenously) and oxygen supplementation.

The next morning our patient showed risus sardonicus and permanent rigor of the upper and lower limbs. These symptoms supported a clinical diagnosis of tetanus. Therefore, he was treated with 10 000 IU of tetanus antitoxin. Antibiotic treatment was changed to metronidazole (20 mg/kg per day). The retrospective tetanus-immunoglobulin G (IgG) level at admission was 2.11 IU/mL.

Twenty-four hours after initiation of the tetanus treatment, his neurologic status remained stable. On day 5 he developed transient bradycardia with prolonged QTc time interval (QT/QTc: 490/476 milliseconds). After 8 days his painful spasms became more infrequent. One day later, he responded to questions by nodding. On day 15, the dysarthria improved, and his speech became partially understandable; on day 17, he showed normal or slightly decreased muscular reflexes and increased muscular tone of all 4 limbs. He was able to sit upright unsupported, and his muscular strength of the upper limbs was 3/5 to 4/5. His right-sided ptosis resolved and improved on the left side. After 3½ weeks the patient was transferred to a rehabilitation center.

Fourteen days after discharge he was able to walk independently with normal power of the upper limbs and lower-limb power at 4/5. His muscular reflexes and fine motor skills were slightly reduced. Minimal ptosis on his left side persisted. At follow-up 6 months and 1 year later his neurologic and psychological examinations were completely normal.

DISCUSSION

Tetanus occurs in different clinical patterns, with generalized tetanus as the most common form. It is caused by the Gram-positive, spore-forming Clostridium tetani, which produces its toxins (tetanospasmin and tetanoylsin) in a favorable environment, preferentially in tissue wounds. Tetanospasmin is able to block neurotransmitter release, which leads to the characteristic increased muscle tone and spasms. In the typical course of tetanus, patients often first notice trismus. Subsequently, dysphagia and stiffness or pain in the upper trunk muscles appears, followed by descending muscular rigidity. Other common clinical manifestations are risus sardonicus, the continuing contractions of face muscles, and an opisthotonos. The clinical course can be complicated by apnea, laryngospasm, aspiration pneumonia, and autonomic dysfunction with need for intensive care management.4

In our patient, initial paralysis and dry mouth, after consumption of grilled meat, misled to a diagnosis of botulism and seemed to be confirmed by the positive mouse toxicity test for botulism: the mice died after injection of the patient’s blood serum but survived after administration of the patient’s serum mixed with botulinum antitoxins A, B, and E (Fig 1). However, as the clinical signs changed from flaccid paralysis to the tetanus-typical rigor, tetanus became clinically obvious despite the patient’s history of appropriate tetanus vaccination. At this stage, the in vivo mouse test had to be doubted. Mice treated with blood serum and botulinum antitoxin, containing antibodies specific for type A, B, or E botulinum toxin separately, became severely paralyzed or died, as did the control mice that received the patient’s blood serum only. The assumed explanation seemed unconventional but simple: single antitoxins originate from rabbits, whereas the combination of A, B, and E botulinum antitoxins is obtained from horses. In contrast to rabbits, horses are routinely immunized against tetanus. Thus, the combined botulinum antitoxin mixture also contained tetanus antitoxin. Conclusive results were obtained by the tetanus mouse test (adapted from the work of Habermann and Wiegand4). Mice that received the patient’s blood serum plus tetanus antitoxin survived without symptoms. This result led to the eventual diagnosis of atypical tetanus in a fully vaccinated child.

On the basis of its exquisite sensitivity, the gold standard of botulinum and tetanus neurotoxin detection is still the mouse toxicity test6,7: the lethal amount for
Botulinum toxin is in the range of 0.5 to 1.2 ng per kg of body weight (depending on the botulinum toxin subtype, intraperitoneal injection route), and for tetanus toxin it is 1 ng per kg of body weight (intraperitoneal injection route). Taking into account the maximal injection volume of 1 mL and an average mouse weight of 20 g, this results in a sensitivity of 10 to 20 pg/mL.

The procedures for the mouse toxicity test for botulinum and tetanus toxin are similar: the patient material is injected intraperitoneally into mice, and symptoms are observed for several hours up to 4 days. In the case of botulinum toxin intoxication, mice sequentially show ruffled fur, labored but not rapid breathing, a characteristic wasp-like abdomen with narrowed waist caused by increased respiratory effort, weakness of limbs that progresses to total paralysis, and gasping for breath followed by death as a result of respiratory failure. In the case of tetanus toxin intoxication, similar symptoms may occur. However, the characteristic wasp-like abdomen with its narrowed waist is only described in mice after administration of botulinum toxin. This symptom is usually missing when testing for tetanus toxin, and spastic paralysis indicates the presence of tetanus toxin.

Death of mice in the absence of neurologic symptoms is not an acceptable indication of botulism or tetanus, because it may be nonspecifically caused by other microorganisms, chemicals present in the test fluids, or injection trauma. Confirmation and exact neurotoxin typing is performed by mouse-protection tests using polyvalent or monovalent neutralizing antibodies (which is better) as in our studies (refs 8 and 9 for botulinum toxin, refs 5 and 12 for tetanus toxin): on simultaneous application of toxin (or patient material) and the respective neutralizing antibodies, the mice are rescued and no symptoms occur.

Currently, the mouse-protection test is still the standard method of choice for quantifying tetanus toxin–neutralizing antitoxin titers. Furthermore, the mouse assay for botulinum toxin is used most frequently for detecting botulinum toxin in foods or patient material or for assessing the potency of the toxin used as a drug in medicine.

In our case, other differential diagnoses such as myasthenia gravis, Guillain-Barré syndrome including variants, encephalitis, lupus erythematosus or other autoimmune reactions, tumor, leukemia, botulism, and intoxication seemed very unlikely, because the results of repeated MRI and laboratory results were completely normal, and the patient’s clinical signs changed quickly from paralysis to rigor. A rare differential diagnosis of tetanus is strychnine poisoning with some similar symptoms such as restlessness, anxiety, muscle twitching, intense pain, trismus, facial grimacing, opisthotonus, and extensor spasm. The rapid onset of symptoms in strychnine poisoning, usually 10 to 20 minutes, made this diagnosis unlikely for our patient, because his clinical picture first showed flaccid paralysis, and rigor of the limbs and risus sardonicus occurred the next day. Hence, a screening for strychnine and alkaloids of Strychnos species was not performed. The intermittently observed bradycardia with prolonged QT-time interval has been described in patients with tetanus.

An increased incidence of tetanus in countries with immunization programs has been reported in elderly adults with impaired immunity despite preceding vaccination. In children with adequate immunization, there have been only a few case reports of tetanus infections. Our patient’s tetanus-IgG level at admission was 2.11 IU/mL, which is considered to be long-lasting protection against infection (range: >1.1 to 3 IU/mL). This level was rechecked at the same laboratory. Unfortunately, no serum was left from the initial blood sample.
for retesting in another institution; the patient had already been treated with immunoglobulin and botulinum antitoxin before the eventual diagnosis of tetanus was made. However, it should be noted that the indicated antitetanus IgG level summarizes protecting and non-protecting antibodies. If the patient has either a low quantity of protecting antibodies in the serum or, alternatively, the concentration of the toxin is too high to be neutralized by the circulating protecting antibodies, the patient develops tetanus and the mouse test for tetanus gives a positive result. Crone and Reder speculated in their case series that burden of toxin can overwhelm patients’ defenses or that an antigenic variability between toxin and toxoid could cause immunization failure.

Treatment of tetanus is based on 3 principles: neutralization of unbound toxin, prevention of additional toxin release, and amelioration of ongoing symptoms. Early, aggressive, intensive care treatment is indicated to prevent or alleviate fatal complications such as respiratory failure and autonomic dysfunction.

Although unintended, but presumably life saving, our patient was treated early for tetanus: he received at least 750 IU of tetanus antitoxin with the botulinum antitoxin (Chiron-Behring GmbH & Co KG, verbal communication, 2005) and an additional 2000 IU with the immunoglobulin infusion (Gamunex 10% has an average content of tetanus antitoxin of 2 IU/mL [Bayer Healthcare AG, verbal communication, 2005]).

CONCLUSIONS

Atypical tetanus should be considered as a rare differential diagnosis in patients with neurologic symptoms despite complete tetanus vaccination. It can be proven unequivocally by the mouse toxicity test.

REFERENCES

Global Climate Change and Children’s Health

Katherine M. Shea, MD, MPH, and the Committee on Environmental Health

ABSTRACT

There is a broad scientific consensus that the global climate is warming, the process is accelerating, and that human activities are very likely (>90% probability) the main cause. This warming will have effects on ecosystems and human health, many of them adverse. Children will experience both the direct and indirect effects of climate change. Actions taken by individuals, communities, businesses, and governments will affect the magnitude and rate of global climate change and resultant health impacts. This technical report reviews the nature of the global problem and anticipated health effects on children and supports the recommendations in the accompanying policy statement on climate change and children’s health.

INTRODUCTION

Scientists and governments concur that Earth is warming: rapid global climate change is underway, and human activities are very likely (>90% probability) the main cause. Adverse human health and ecosystem consequences are anticipated, and some are already being measured. Physicians have written on the projected effects of climate change on public health, but little has been written specifically about anticipated effects of climate change on children’s health.

Children represent a particularly vulnerable group that is likely to suffer disproportionately from both direct and indirect adverse health effects of climate change. Pediatric health care professionals must understand the escalating nature of these threats, anticipate their effects on children’s health, and participate as children’s advocates for strong mitigation and adaptation strategies now and at all levels, from local to global. This technical report examines both direct and indirect threats to children’s health and futures related to climate change.

NATURE OF THE GLOBAL PROBLEM

“Warming of the climate system is unequivocal, as is now evident from observations of increases in global average air and ocean temperature, widespread melting of snow and ice, and rising global mean sea level.” According to the National Climatic Data Center, all records indicate that during the past century, global surface temperatures have increased at a rate near 0.6°C per century (1.1°F per century), but the trend has been 3 times larger since 1976. The results of this warming on regional climate are not uniform. In general, land-surface temperatures are increasing faster than sea-surface temperatures. The climate in latitudes between 40°N and 70°N is warming more quickly than that in lower latitudes, and some areas (eg, the southeastern United States) are actually cooling. Changes in precipitation that occur with climate change...
are also nonuniform. Since 1900, precipitation has increased 5% globally, but it has increased 0.5% to 1% per decade in northern midlatitudes and decreased 0.3% per decade in subtropical latitudes. In contrast, snowfall in the northern hemisphere has decreased by 10% since 1966.

Examples of the effects of climate change have been widely reported. Glaciers are in rapid retreat, and Arctic sea ice is melting. As a result of thermal expansion, sea level has increased 1 to 2 mm/year over the past 100 years. Oceans are acidifying as atmospheric carbon dioxide (CO$_2$) is absorbed by the marine buffer system. Ecosystems and individual species are being affected in a variety of ways. Changes in temperature affect the density and range of species; natural history traits such as migration, flowering, and egg laying; morphology such as body size and behavior; and genetic frequency shifts. In an analysis of 143 studies that span decades of observation, more than 80% of 1468 species (mollusks to mammals and grasses to trees) are currently showing significant changes in temperature-sensitive species traits.

There is strong consensus among expert scientists that Earth is undergoing rapid, global climate change, although there remains uncertainty about how rapidly and extensively the climate will change in the future. Given the range of possibilities, the Intergovernmental Panel on Climate Change has developed a suite of scenarios for different levels of mitigation and adaptation in response to anthropogenic (man-made) global climate change; all their cases predict that temperatures and sea level will continue to rise throughout the 21st century. Recent analyses describe thermal inertia in Earth’s climate system such that even if greenhouse gas (GHG) emissions were abruptly reduced to zero, the planet would continue to warm for decades until the energy stored in the system equilibrates. The possibility of reaching a tipping point at which abrupt, large, and irreversible change could be superimposed on current trends adds both urgency and further ambiguity to the situation. In this context, it is critical to understand that current human activities are accelerating climate change and that future human activities will affect their trajectories.

ANTHROPOGENIC CAUSES OF THE CHANGE

The greenhouse effect is necessary to life on Earth as we know it (Fig 1). Without heat-trapping GHGs such as water vapor, CO$_2$, and other natural components of the atmosphere, Earth would be a lifeless, frozen planet (average temperature: $-18^\circ$C) instead of the diverse biosphere we know today. Since the onset of the industrial age, however, human activity has dramatically enhanced the greenhouse effect by rapidly adding large amounts of GHGs to the atmosphere (Table 1 [note that the United States leads total country and per-capita emissions]). Three GHGs, CO$_2$, methane, and nitrous oxide, are responsible for approximately 88% of the anthropogenic influences that enhance the greenhouse effect and have increased 35%, 155%, and 18%, respectively, since 1750 (the beginning of the industrial era). Rates of increase in GHGs are accelerating, up 20% since 1990.

CO$_2$ is the most important GHG and is responsible for more than 60% of human-enhanced increases and more than 90% of rapid increase in the past decade. Most CO$_2$ emissions are from the burning of fossil fuels such as coal, oil, and gas. Rising CO$_2$ is also related, to a lesser extent, to deforestation, which eliminates an important carbon sink (carbon sinks are reservoirs that absorb or take up released carbon from another part of the carbon cycle; the 4 major sinks on the planet are the atmosphere, oceans, terrestrial plants, and soils).
sphere, the terrestrial biosphere [eg, trees and freshwater systems], oceans, and sediments. Currently, the atmosphere contains approximately 370 ppm of CO₂, which is the highest concentration in 420,000 years and perhaps as long as 2 million years. Estimates of CO₂ concentrations at the end of the 21st century range from 490 to 1260 ppm, or a 75% to 350% increase above preindustrial concentrations.

The importance of the magnitude of GHG emissions is linked to the rate of release. In the distant geologic past, similar concentrations of atmospheric CO₂ have occurred, but they accumulated over a 10,000-year period, allowing for the slow, global biogeochemical cycles to adjust to the increases. Current emissions are being added to the atmosphere at 300 times this rate. This confluence of speed and quantity of emissions has created the current, unprecedented rapid climate change.

CLIMATE CHANGE–ASSOCIATED HEALTH EFFECTS ON CHILDREN
Human health is affected by the condition of the physical environment. Because of their physical, physiologic, and cognitive immaturity, children are often most vulnerable to adverse health effects from environmental hazards. As the climate changes, environmental hazards will change and often increase, and children are likely to suffer disproportionately from these changes. Anticipated health threats from climate change include extreme weather events and weather disasters, increases in certain infectious diseases, air pollution, and thermal stress. Within all of these categories, children have increased vulnerability compared with other groups. These direct health threats are discussed in this section, with an emphasis on children in the United States. Indirect threats are discussed briefly in “Long-term and Indirect Climate Change-Associated Health Threats to Children” below.

Extreme Weather Events and Weather Disasters
The Intergovernmental Panel on Climate Change predicts that it is “likely” or “extremely likely” that climate change will cause increased frequency and intensity of extreme weather events and weather disasters. Often, these events are categorized as floods, storms, and droughts. Floods represented 43% of weather-related disasters between 1992 and 2001 and are the most frequent weather-related disaster. Although less prevalent, droughts and their associated famines are the most deadly weather-related disasters. Developed countries such as the United States have systematically increased the risk to populations from flood events by developing coastlines and flood plains. In the United States, hurricanes and tornadoes may be the most dramatic and visible weather disasters. Evidence suggests that the frequency of category 4 and 5 hurricanes has increased over the past 30 years, but the observation period is still too short to attribute this change to increased sea-surface temperature and climate change with high confidence. The health consequences associated with extreme weather events include death, injury, increases in infectious diseases, and posttraumatic mental health and behavior problems. Few studies have specifically examined such consequences in children. Globally, 66.5 million children annually were affected by disasters between 1990 and 2000. Children everywhere are at risk of injury and death from storms and floods. In the developed world, infectious disease outbreaks follow natural disasters when sanitation, sewage treatment, and water-purification plants become damaged or overwhelmed, refrigeration and cooking facilities are disrupted, and people are unusually crowded in temporary shelter. These outbreaks are usually mild and well controlled, which is in contrast to the aftermath of similar catastrophes in developing nations, where disease outbreaks can be deadly. Mosquito-borne and other vector-borne illnesses may also be increased when storms or floods create large amounts of standing water suitable for breeding. Mental and emotional distress documented for children and adolescents after weather disasters include posttraumatic stress disorder and high rates of sleep disturbance, aggressive behavior, sadness, and substance use/abuse. Some studies have suggested that children have more persistent symptoms than adults who experience the same disaster, but more studies specific to children’s experience are required. Community support services and early therapeutic intervention and postdisaster counseling can significantly reduce the medium- and long-term mental health burden on children. Experiences with Hurricane Katrina demonstrated the difficulties with tracking children’s whereabouts, keeping children and caregivers together, and special needs of hospitalized infants and children during and after major natural disasters.

Infectious Diseases
Globally, infectious diarrhea is the second-leading cause of death in young children; water-borne gastroenteritis is projected to increase under conditions of global warming. Currently, the World Health Organization estimates that, approximately 1.62 million children younger than 5 years die of diarrhea annually, and most cases are attributable to contaminated water. Although children in developed countries are unlikely to die of water-borne infections, they may suffer illness that is attributable indirectly to climate change. Events associated with El Nino serve as a model for global warming by altering weather for periods of several years in the direction of a hotter climate. During El Nino events, rates of hospitalizations of children for diarrhea increase. (In 1 study, the rate of hospitalizations of children for diarrhea increased 8% per degree centigrade of temperature increase.) Water-borne disease outbreaks in the United
Vector-borne infections are affected by climate change.42 Both the hosts (eg, rodents, insects, snails) and the pathogens (eg, bacteria, viruses, parasites) can be sensitive to climatic variables such as temperature, humidity, and rainfall. The ability to predict disease rates related to climate change is complicated by a large number of additional variables such as topography, land use, urbanization, human population distribution, level of economic development, and public health infrastructure.43 There is no easy formula that predicts climate change–related infection risk with confidence.

Malaria is a climate-sensitive vector-borne illness to which children are particularly vulnerable. According to the World Health Organization, malaria currently causes than 1 million deaths.44 Because they lack specific immunity, children experience disproportionately high levels of both morbidity and mortality from malaria; 75% of malaria deaths occur in children younger than 5 years. The young are also more susceptible to cerebral malaria, which can lead to lifelong neurologic damage in those who survive. In areas of sub-Saharan Africa, the death rate from malaria in children 0 to 4 years of age is 9.4 in 1000 vs 0.13 in 1000 in those older than 14 years.45 More than 3 billion people live in malaria-prone areas today. Climate change is expanding the range of host mosquitoes to higher altitudes and higher latitudes, and warmer temperatures speed the development of the parasite within the host vector.46 Small children will be most affected by the expansion of malaria zones and the success or failure of societal response to this change.

Three vector-borne diseases that affect the United States illustrate ways in which climate change can enhance disease burden: West Nile virus infection, Lyme disease, and hantavirus pulmonary syndrome.

West Nile virus infection was first reported in the United States in New York in 1999. Although it is still not known how it entered the United States, once introduced, it spread rapidly. A series of warm winters failed to kill the mosquito vectors. Warmer summers amplified the life cycle of the mosquitoes and increased the viral load. Drought and rain cycles, particularly as they affected urban landscapes, increased the contact of the bridging mosquito vectors with birds and humans.46 Human populations with no herd immunity were highly susceptible to infection. In 1999, there were 62 human cases of West Nile virus infection, all reported from New York state. In 2003, there were 9862 human cases reported from 45 states and the District of Colum-
communities with high levels of particulate air pollution.55

A second change that is being observed is the temperature-related increases in pollen production and other aeroallergens in some regions and some cities. Increased temperature causes increases in amounts of pollens produced by some plants56 and can also affect spatial distribution and density of plants, fungi, and molds that produce aeroallergens.60 To the extent that exposure to aeroallergens contributes to the incidence, prevalence, and severity of asthma, atopy, and other respiratory disease, climate change will affect the pattern of disease in children. Some investigators have argued that part of the current global increase in childhood asthma can be explained by increased exposure to aeroallergens driven by climate change.61

**Thermal Stress**

For all organisms, there exists a range of ideal temperature above and below which mortality increases. Humans are no exception, although temperature-mortality relationships vary significantly by latitude, climatic zone, and level of socioeconomic development.3 As ambient temperatures increase, the frequency of heat waves will increase. It is expected that there will be fewer cold-related deaths in a warmer world,62 but whether this will offset the expected increase in heat-related deaths is unknown. Populations that live in temperate climates, such as in the United States and Europe, are likely to be hard hit initially, because global warming is most dramatic in these latitudes and there has been little time for populations to acclimatize to changes in temperature. Observations on heat and mortality have been reported for decades63 and have gained recent attention with the heat waves of 2003 in Europe64 and of 2006 in Europe and North America.63,65 Heat-related deaths and hospitalizations are most common in the elderly, especially if they are ill.66,67 One study has found that infants and young children may represent a second, albeit smaller, higher-risk group,68 but effects on children have not been studied adequately. In addition, children spend more time outside, especially playing sports in the heat of the afternoon, which puts them at increased risk of heat stroke and heat exhaustion.69 Increased outdoor time during hot weather may also put children at increased risk of UV radiation–related skin damage, including basal cell carcinoma and malignant melanoma.70 Some data indicate that heat-related mortality in the United States has decreased in recent years, in part associated with increasing percentage of homes with air conditioners.71 It is currently unknown how effective adaptation and acclimatization will be in preventing excess heat-related deaths and illness.72,73

**LONG-TERM AND INDIRECT CLIMATE CHANGE-ASSOCIATED HEALTH THREATS TO CHILDREN**

Long-term and indirect effects on children’s health from climate change will depend on how the climate continues to change over the next decades and what sorts of mitigation and adaptation strategies are adopted now.17 How quickly and comprehensively GHG emissions can be stabilized and then reduced will have a significant effect on the rate and degree of warming, but even the most optimistic scenarios describe continued warming through the end of this century.17 Food availability may be affected as land and ocean food-productivity patterns shift.79 Water availability may change and become much reduced in some regions, including during summer in the snow run-off–dependent American west coast.75 Coastal populations will be forced to move because of rises in sea level, and massive forced migrations, driven by abrupt climate change, natural disaster, or political instability over resource availability, are conceivable.24 In addition, world population is expected to grow by 50% to 9 billion by 2050, which would place additional stress on ecosystem services and increase the demand for energy, fresh water, and food.54 As these changes evolve, social and political institutions will need to respond with aggressive mitigation strategies and flexible adaptation strategies to preserve and protect public health, particularly for children.

**MITIGATION AND ADAPTATION STRATEGIES**

Strategies to address the effects of climate change, known as mitigation and adaptation, are concepts that parallel the focus on both primary and secondary prevention strategies in pediatric health care. These strategies are discussed briefly here. The prevention or minimization of the effects of climate change on children’s health is beyond the control of an individual pediatrician. Yet, pediatricians can play important public roles as advocates by individual example and through community participation, political involvement, or collective advocacy at the local, state, and national levels.76,77

Broadly, mitigation policies (Table 2) for reduction of atmospheric GHG include reducing emissions through energy efficiency and use of renewable energy sources, increasing carbon sinks by forest preservation and reforestation, and development of GHG-capture and -storage technologies (carbon sequestration is the fixation of atmospheric CO₂ in a carbon sink through an active process). Adaptation involves developing public health strategies to minimize adverse health outcomes that are anticipated from climate change. These strategies include improved disease surveillance and reporting, improved weather forecasting and early warning systems, advanced emergency management and disaster-preparedness programs, development and dissemination of appropriate vaccines and medicines, and public health education and preparedness. Category-specific examples
can be found at www.grida.no/climate/ipcc.tar/ wg2/646.htm#tab18-2. These adaptation strategies include policy and legislative actions, engineering responses, and personal behavior change.

Effective implementation of mitigation and adaptation strategies must involve actions from the global to local levels by governments, corporations, communities, and individuals. Furthermore, climate change is part of generalized global change, which includes population growth, land use, economic change, and evolving technology; all have effects on individual human and public health (Fig 2). Any solutions that address climate change must be developed within the context of overall sustainable development (the use of resources by the current generation to meet current needs while ensuring that future generations will be able to meet their needs). Protecting the health of current and future generations requires a fundamental shift in thinking for health professionals: pediatricians, as advocates for children’s health, can be leaders in a move away from a traditional focus on disease prevention to a broader, more integrated focus that encompasses sustainability as synonymous with health. Given the health implications for current and future generations of children, the disease-prevention role for pediatric health care professionals includes advocating for environmental sustainability.

**SUMMARY**

This technical report describes the broad scientific consensus that man-made climate change has begun...
and is accelerating. The major cause of this change is the rapid release of CO₂ from burning of fossil fuel. All predictions indicate that climate change will continue for at least a century, but the trajectory of that change depends on human responses. There are anticipated effects on human health from extreme weather events, infectious diseases, air pollution, and heat stress. Although little research thus far has concentrated on the pediatric age group, it is likely that children will suffer disproportionately from climate change. Furthermore, the state of the world of future children is uncertain and depends on actions taken to mitigate and adapt to climate change and other global-scale trends. Pediatric health care professionals are in an ideal position to advocate for action, not only to address climate change but also, more broadly, to ensure sustainability. Specific recommendations for pediatricians and governments are enumerated in the American Academy of Pediatrics policy statement on climate change and children’s health, which accompanies this technical report.

COMMITTEE ON ENVIRONMENTAL HEALTH, 2006 –2007
Michael W. Shannon, MD, MPH, Chairperson
Dana Best, MD, MPH
Helen J. Binns, MD, MPH
Joel A. Forman, MD
Christine L. Johnson, MD
Catherine J. Karr, MD, PhD
Janice J. Kim, MD, PhD, MPH
Lynnette J. Mazur, MD, MPH
James R. Roberts, MD, MPH

PAST COMMITTEE MEMBER
Katherine M. Shea, MD, MPH

REFERENCES
10. Dore MH. Climate change and changes in global precipitation patterns: what do we know? Environ Int. 2005;31:1167–1181
16. US Environmental Protection Agency. Climate change-science:
52. Centers for Disease Control and Prevention. Seal up! Trap up!


TECHNICAL REPORT

Ventricular Fibrillation and the Use of Automated External Defibrillators on Children

David Markenson, MD, Lee Pyles, MD, Steve Neish, MD, and the Committee on Pediatric Emergency Medicine and Section on Cardiology and Cardiac Surgery

ABSTRACT
The use of automated external defibrillators (AEDs) has been advocated in recent years as a part of the chain of survival to improve outcomes for adult cardiac arrest victims. When AEDs first entered the market, they were not tested for pediatric usage and rhythm interpretation. In addition, the presumption was that children do not experience ventricular fibrillation, so they would not benefit from use of AEDs. Recent literature has shown that children do experience ventricular fibrillation, and this rhythm has a better outcome than do other cardiac arrest rhythms. At the same time, the arrhythmia software on AEDs has become more extensive and validated for children, and attenuation devices have become available to downregulate the energy delivered by AEDs to allow their use in children. Pediatricians are now being asked whether AED programs should be implemented, and where they are being implemented, pediatricians are being asked to provide guidance on the use of AEDs in children. As AED programs expand, pediatricians must advocate on behalf of children so that their needs are accounted for in these programs. For pediatricians to be able to provide guidance and ensure that children are included in AED programs, it is important for pediatricians to know how AEDs work, be up-to-date on the literature regarding pediatric fibrillation and energy delivery, and understand the role of AEDs as life-saving interventions for children.

INTRODUCTION
Early defibrillation has been shown to be the most effective treatment for adult out-of-hospital cardiac arrest caused by ventricular fibrillation (VF). The likelihood of survival decreases by approximately 7% to 10% with each minute of delay to defibrillation after cardiac arrest. Strategies to decrease the time to defibrillation that have been shown to be effective include the use of an automated external defibrillator (AED) by emergency medical services (EMS) personnel and nonmedical lay people. AEDs represent a significant breakthrough for adult out-of-hospital cardiac arrest. For adults, when combined with effective cardiopulmonary resuscitation (CPR), the use of early defibrillation has been shown to produce the highest rates of survival.

For children, the use of defibrillation traditionally has been downplayed with a focus on early airway and ventilatory assistance, because available data showed that asystole was the predominant rhythm and that VF rarely occurred. Although not the most common rhythm, VF does occur in children. In addition, the chance of surviving after VF is greater than that for other nonperfusing rhythms, which
makes timely treatment of VF a priority for pediatric resuscitation. As a result, many current guidelines, including those of the International Liaison Committee on Resuscitation, the American Heart Association, and the National Association of EMS Physicians, now advocate for AED use on children to analyze rhythms and provide early defibrillation in cases of VF.

To be effective and safe for children, an AED must achieve several goals. First, it must capture the patient’s rhythm from surface electrodes and then use a computer algorithm to determine if a shock is indicated. For an AED to be used on children, it must have the capability to determine shockable rhythms, but more importantly, it must accurately determine when not to deliver a shock. When the correct rhythm is identified, the AED must be capable of delivering a shock with sufficient energy to convert the rhythm to a perfusing rhythm without causing damage to the myocardium. In the past, published data related to the effective energy for defibrillation or safety of different energy levels and waveforms for children were lacking. In previous reviews, the only source of data regarding energy levels was a single retrospective study from 1976, which showed, on average, that 2 J/kg treated VF in many children and, if it failed, that 4 J/kg was usually effective. As a result, AEDs were recommended for use only on older children and adults. Fortunately, AED technology has improved dramatically in recent years, and additional research has been conducted and demonstrated that AEDs are safe for use on younger children. In addition, although there are fewer data available for infants and young children than for adults and older children, there are studies that have shown AED safety and efficacy for infants and young children. Finally, pediatric-capable AEDs have been approved by the US Food and Drug Administration not only for use on young children but also on infants of all ages. AEDs are now capable of recognizing pediatric shockable rhythms, and some are programmed to decrease the delivered energy on the basis of a fixed reduction, chest wall impedance, or a combination of the two, which makes them suitable for pediatric patients.

CARDIAC ARREST PHYSIOLOGY
Children and adults have anatomic differences that may be significant in pediatric defibrillation. Children’s hearts are smaller than those of adults. A critical mass of myocardial tissue is required to sustain fibrillation. This could be one of the reasons why VF is less prevalent in the pediatric population than it is in adults.

Pediatric and adult patients also have some important physiologic differences. Children have higher cardiac output per kilogram of body weight than adults, but because oxygen demand is high in children, oxygen reserves are limited. Cardiopulmonary deterioration can occur whenever oxygen delivery is compromised or when oxygen demand is increased above oxygen supply. Children have higher heart rates and lower stroke volumes than adult patients.

In children, sinus tachycardia is the normal response to stress, because infants and children increase their cardiac output by increasing their heart rate rather than stroke volume. Normal heart rates for neonates have been reported to range from 100 to 180 beats per minute. In addition, detecting a carotid pulse in infants may be more difficult than it is in adults because of their shorter, chubbier necks.

Adults and children also have biochemical differences, which may be relevant to the toxicity of defibrillation shocks. Newborn infants have substantially less myocardial catecholamine than adults. It is believed that the biochemical effects of catecholamines on oxygen consumption and use may have a role in causing myocardial damage. Thus, it may be expected that newborn infants would have a higher tolerance for high-energy defibrillation doses than adults.

CARDIAC ARREST EPIDEMIOLOGY
Children suffer fewer cardiac emergencies than adults. One estimate indicates that sudden cardiac death is one tenth as common in children as it is in adults and that it occurs in only 1 to 2 per 100 000 children annually. However, the death of a child is an enormous emotional and social loss and has a community-wide impact. Because of their life expectancy, the number of years of life lost as a result of pediatric cardiac arrests may rival that for all adult arrests. Survival and neurologic outcomes are better for pediatric patients whose initial recorded rhythm is VF, compared with other causes of pediatric cardiac arrest.

In a 6-year retrospective population-based review of pulseless, nonbreathing patients younger than 20 years, Mogayzel et al compared the causes and outcomes for patients whose initial rhythms were VF with those whose initial rhythms were asystole or pulseless electrical activity (PEA). Of the 157 patients included in the study, VF was the initial rhythm of 19%, excluding patients younger than 6 months who died as a result of sudden infant death syndrome. The study reported that, in the witnessed arrests, a significant percentage of the patients were found to be in VF as the initial rhythm on arrival of EMS. In addition, it was reported that the age of the patient did not alter the percentage of patients with VF. The percentage of cardiac arrest patients found in VF was approximately the same for 0- to 4-year-olds as it was for 15- to 19-year-olds (17% vs 19%, respectively). The first responders identified the initial rhythm of only 44% of the patients. A majority of the patients in this study (16 of 29 children) who initially were in VF did not receive defibrillation by the first responder, because protocol and available equipment required the proce-
and young children, AEDs that advise shocks primarily on the basis of heart rate would not be appropriate for use with pediatric patients. This is of even greater importance for a public-access defibrillation program in that most lay people are taught not to check the patient’s pulse before application of an AED but to rely on the device’s recognition of a shockable rhythm. The anatomic and physiologic differences between adults and children summarized previously underscore the importance of this requirement. Because many high-rate rhythms in children may be associated with a pulse, it is important that an AED for use on children be designed so that shocks are not advised for such rhythms. For rhythms for which there is not general agreement regarding whether a shock is warranted and for which there may be an associated pulse (intermediate rhythms), the AED should be designed such that shocks are not advised. The American Heart Association defines intermediate rhythms as “rhythms for which the benefits of defibrillation are limited or uncertain.”

For these rhythms, the therapeutic benefit of a shock is uncertain, and the victim may be exposed to some risk if a shock is delivered.

Studies of AED rhythm detection in children generally have reported good accuracy. In 1 study, rhythms from patients in PICUs with ages ranging from 5 days to 7.5 years were recorded. Rhythms were digitized and annotated by 3 reviewers and then read into several AEDs. The AED-analysis results were then compared with those of the reviewers. The study found that the specificity of the AEDs for pediatric tachycardias was not 100%, which suggests that modifications to AED algorithms might be needed to accommodate pediatric patients.

Atkins et al compared rhythms obtained by using an AED with those obtained by using a standard electrocardiography device and then digitally analyzed both the sensitivity and specificity of the AED software in identifying pediatric VF. In this study, 696 five-second rhythm strips were analyzed. The AED was found to have a specificity of 100% for nonshockable rhythms and sensitivity of 96% for detecting VF. The authors demonstrated that AEDs have a low risk of providing an inappropriate shock and that the AED correctly identified shockable rhythms, which makes their software algorithm safe and effective for children.

In 1 of the only studies of out-of-hospital AED use on children younger than 16 years, AEDs were found to be highly accurate, with 100% specificity and 88% sensitivity. The AEDs used in this study included multiple parameters in their analysis algorithms (the criteria on which these AEDs will advise a shock includes more
than just heart rate and amplitude). Although sensitivity was excellent at 88%, this result does present some missed opportunities and room for improvement. More important, however, was the 100% specificity, which indicates that under no circumstances was a patient defibrillated incorrectly. The authors of an AED manufacturer’s study of its patient-analysis system collected 233 rhythm strips from 71 pediatric patients younger than 12 years and showed similar results, with 100% specificity and sensitivity for VF and 81.8% for VT. The study supported the feasibility of using a single AED algorithm for both adults and children.

In a case report that described the use of an AED on a patient younger than 8 years, the AED correctly detected VF and advised a shock, which defibrillated the VF successfully, and then the device correctly detected the resulting nonshockable rhythm and advised that no shock was required. Importantly, there was no detectable cardiac damage from the defibrillation.27

A recent observational study by Atkins and Jorgenson28 of pediatric-attenuated pads used to reduce the energy delivered by some AEDs showed that for patients younger than 10 years, the 8 VF rhythms were identified correctly, as were the 10 nonshockable rhythms.

**AED ENERGY AND WAVEFORM SAFETY AND EFFICACY**

Defibrillation involves delivery of current through the chest and to the heart to depolarize myocardial cells and eliminate VF. In addition to recognizing a shockable rhythm, an AED must deliver the electrical energy that will have the greatest potential for conversion to a perfusing rhythm while minimizing the potential for harm. The currently accepted prehospital approach for manual defibrillation is 2 J/kg initially, followed by 4 J/kg. The energy range for current AEDs is between 150 and 360 J, depending on type of waveform and model.

Modern defibrillators are classified according to 2 types of waveforms: monophasic and biphasic. Monophasic waveform defibrillators were introduced first, but biphasic waveforms are used in almost all the AEDs and manual defibrillators that are available today. Energy levels vary according to type of device. No specific waveform (either monophasic or biphasic) is consistently associated with a higher rate of return of spontaneous circulation or improved rates of survival to hospital discharge after cardiac arrest. Monophasic waveforms deliver current of 1 polarity. Monophasic waveforms can be categorized further by the rate at which the current pulse decreases to 0. The monophasic damped sinusoidal waveform returns to zero gradually, whereas the monophasic truncated exponential waveform current is abruptly returned to baseline 0 current flow. Few monophasic waveform defibrillators are being manufactured, but many are still in use. Most of these defibrillators use monophasic damped sinusoidal waveforms. Biphasic AED models primarily come in 2 types of waveforms: truncated exponential and rectilinear. Both of these waveforms are fairly prevalent in available biphasic defibrillators.

In 1976, Gutgesell et al29 reported the results of a retrospective chart review of pediatric cardiac arrests that was performed to determine if their institution’s guidelines for defibrillation energy dose in pediatric patients was effective. Twenty-seven children were included, with weights ranging from 2.1 to 50.0 kg and ages ranging from 3 days to 15 years. In these children, 71 defibrillation attempts were made by using monophasic wave technology. The authors found that 91% of the shocks with an energy dose of 2 J/kg ± 10 J were effective, whereas the 2 shocks below this level were ineffective, and all but 1 of the 12 shocks above this level were effective. The single higher-energy shock that was ineffective was only 13 J higher than the 2 J/kg dosage guideline, and the child previously had received an unsuccessful initial shock at a lower level (1.9 J/kg). A third shock at 60 J (3.8 J/kg) was successful for this patient. All shocks above 2 J/kg were at least as effective as shocks of 2 J/kg ± 10 J, and the study included 1 shock at more than 7 J/kg.30

Before the study reported by Gutgesell et al, the recommendations for initial defibrillation energies of children ranged from 60 to 200 J.30,31 Afterward, this study was the basis for the pediatric defibrillation recommendations of 2 J/kg, followed by 4 J/kg, for children. It is important to note that the study by Gutgesell et al was not designed to establish a defibrillation threshold (DFT)–based dose for children; rather, it confirmed the guidelines for effective defibrillation that the investigators had already established on the basis of their previous animal studies. In addition, this study established neither dose safety ranges nor the risk/benefit of this or any other dosing strategy. Although Gutgesell et al acknowledged that the damage threshold is much higher, they decided to keep the 2 J/kg initial and 4 J/kg subsequent shocks protocol, because it was easy to remember and seemed to be successful in most cases. This single retrospective case series, in which a monophasic waveform was used, has served as the basis for most guidelines for pediatric defibrillation, including extrapolation without evidence to biphasic waveform energy settings.

Mogayzel et al found similar results, reporting that 93% of the 29 patients in their study whose initial rhythm was VF were defibrillated with 2 to 4 J/kg as the initial dose and 4 J/kg for subsequent shocks. Seventeen percent of the patients with VF were discharged with no or mild disability, compared with only 2% of patients whose initial rhythm was asystole or PEA. The authors found that VF was the only variable associated with a good outcome in this population and that patients whose initial rhythms were VF had a survival rate that approached that of adults.

Atkins et al recently studied out-of-hospital AED use on patients younger than 16 years. In total, the AEDs...
performed 67 analyses, and there were 25 episodes of VF. Three of the 7 patients who received shocks survived and were discharged from the hospital, with AEDs having delivered either 200 or 360 J. Two of the patients with VF did not receive shocks, because they were younger than 12 years and/or weighed less than 90 kg, and the dose would have exceeded the 4 J/kg recommended dose; both patients died.

Successful use of an AED in a younger child has been reported. In this case, a 3-year-old child received a 9 J/kg dose delivered by his mother using a 150-J biphasic AED. The child was awake and crying with a heart rate of 120 beats per minute when the EMS team arrived 10 minutes after his defibrillation. The child was defibrillated successfully with 1 shock. The child’s creatinine kinase and troponin concentrations were within the reference ranges after his resuscitation, which indicated no clinically significant damage. A postmarket observational study of pediatric-attenuated pads by Atkins and Jorgenson has shown that using these pads with a certain AED and a fixed biphasic energy of 50 J resulted in successful termination of 8 cases of VF.

The harmful effects of defibrillation cannot ethically be induced deliberately in human subjects; therefore, it is not possible to determine the dose-response curves for toxic and lethal energies in humans. As a result, appropriate energy dosing for defibrillation and energy dosing that might cause injury must be extrapolated from animal models.

Determination of the DFT for a waveform typically requires delivery of multiple shocks to each subject in a controlled setting, which makes the study in a human population extremely difficult; the difficulty of studying effectiveness of external defibrillations shocks in human pediatric subjects is compounded by the lower incidence of pediatric cardiac arrest. One study across a wide variety of animal species showed that the energy dose required for defibrillation was somewhat weight dependent and ranges from 0.5 to 10 J/kg. Although external defibrillation-effectiveness data are too limited to draw any definitive conclusions, this study suggests that the DFT for most patients who weigh up to 50 kg is likely close to 2 J/kg. Other studies in animals and humans have shown that repeated high-energy shocks with a 360-J monophasic damped sine waveform might cause significant damage. One study found that animals that received a single high-energy shock sustained little damage, but animals that received multiple shocks had significant cardiac injury and acute pump failure. The authors emphasized the need to optimize first-shock effectiveness.

A clinical study also has shown that initial shocks that were too low (below the DFT) caused an increase in the energy requirement for subsequent shocks to defibrillate. This report emphasizes the importance of first-shock effectiveness and casts doubt on the philosophy of starting with a low or moderate energy level with the intent to increase if needed for subsequent shocks. These reports also provide a possible explanation for the single unsuccessful shock that was more than 2 J/kg in the pediatric patient in the Gutgesell et al study. This patient initially received a shock that was less than 2 J/kg. In addition, these data suggest that an initial dose somewhat higher than the current recommendations may be warranted for pediatric patients.

In 1980, Babbs et al published therapeutic indices for effective, damaging, and lethal doses of defibrillation energy using monophasic wave technology on the basis of studies that involved more than 100 dogs. The delivered energies ranged from 1 to 512 J/kg. They found that it took 5 times more energy to produce detectable histologic damage than was required for effective defibrillation. The ED50 (defined as the energy at which 50% of the animals were defibrillated successfully) was 1.5 J/kg. The TD50 (defined as the energy at which 50% of the animals had detectable myocardial damage) was 30 J/kg. The LD50 (the 50% lethal dose for the population) was 470 J/kg. The ED50 curve was significantly steeper than the TD50 and LD50 curves, showing that the toxic and lethal effects were more variable in the study population. The authors concluded that the “fear of inducing damage should not be a dominant factor in determining defibrillation dose. Instead, effectiveness should be the major criterion.”

Gaba and Talner studied monophasic wave defibrillation safety in 21 newborn pigs that ranged from 2 to 18 days of age and weighed 0.95 to 4.7 kg. Some animals that were shocked with doses greater than 150 J/kg had substantial myocardial damage, but this effect was not seen in animals that were shocked at lower energy levels. The dose-response curves established by this study are nonlinear and exponential, as seen in other studies. However, the authors noted that substantially more energy was needed to cause myocardial damage in newborn piglets when compared with results reported in previous investigations of adult dogs. These results suggest that in humans, newborns are more tolerant of high-energy doses than adults, and Gaba and Talner hypothesized that the intrinsic structural and physiologic differences between the newborn and adult myocardium could account for the differences observed.

Several studies have shown that biphasic waveforms cause less damage than monophasic waveforms, and these effects seem to be independent of energy. One study found that the use of biphasic waveforms was associated with significantly better postresuscitation myocardial function than monophasic waveforms, even with the same high energies and capacitance typically used for monophasic defibrillation. These studies demonstrated the safety of pediatric defibrillation shocks, even at energy doses significantly higher than the cur-
rently recommended 4 J/kg and especially when the number of shocks delivered is minimized.

Several recent studies have addressed the question of energy dosage by using a biphasic AED and the benefit of biphasic versus monophasic waveforms. In a study of piglets, Clark et al. showed that biphasic waveforms proved superior to monophasic waveforms and at lower energies in both infant and young-animal models. Although the exact energy needed for humans cannot be extrapolated from this study, the benefit of biphasic waveforms was shown. In 2002, Killingsworth et al. published results from a piglet study that was aimed at determining the DFT for biphasic truncated exponential waveform shocks. They found that the DFT with pediatric patches was 2.4 ± 0.81 vs 2.1 ± 0.65 J using adult patches. Also of note, initially after the shock, they found an increased drop in left ventricular (LV) function with increasing energy dosage compared with baseline, but by 60 seconds, there was no difference in LV function with increasing energy dosage. The importance of this study was that it determined a 2.3 J/kg DFT but also that higher dosages, even up to 360 J, produced only transient ST segment and hemodynamic changes. This led the authors to conclude that AED dosages of 50 to 100 J would be appropriate for a child weighing up to 25 kg, but using an AED with a higher energy would not pose a risk to a child or even an infant.

Additional studies have evaluated the use of adult AEDs on animal models of children and have compared attenuated adult AEDs versus nonattenuated adult AEDs. A study by Berg et al. in 2003 compared a single escalating energy sequence (50, 75, and 86 J) of an attenuated adult-dose biphasic shock, nonattenuated adult-dose biphasic shock, and weight-based monophasic shock for piglets with prolonged VF. The attenuated and monophasic dosages were delivered via pediatric pads and the nonattenuated dose was delivered via adult pads. The best outcome in terms of 24-hour survival and neurologic function, as well as a lesser decrease in LV end-diastolic function, was found with the attenuated adult-dosage AED. In addition, the study indicated that the adult biphasic AED, although not as good as the attenuated dosage, was superior in outcome and LV ejection fraction to the weight-based monophasic AED. In another study, Berg et al. showed that escalating attenuated adult-dosage biphasic shocks are more effective and have fewer adverse outcomes than standard escalating adult-dosage biphasic shocks. Although the attenuated dosage was more effective and had fewer adverse effects on the myocardium, the adult escalating-dosage biphasic shocks also were effective in terminating VF in this piglet model of pediatric VF. Berg et al. have also shown that attenuated adult-dosage biphasic shocks are at least as effective, if not more effective and safe, than weight-based monophasic shocks in a piglet model. A study by Tang et al. showed that a single attenuated adult-dosage biphasic energy dosage (50 J) was effective in terminating VF in piglets that ranged in weight from 3.8 to 25.0 kg. In addition to the effective termination of VF, even when applied to piglets that weighed only 3.8 kg (comparable with an infant), this energy dosage resulted in a return to normal for both hemodynamic and myocardial function, indicating the safety of the attenuated adult-dosage technique, even in models of newborn infants.

The animal data indicate that there is a wide margin between effective and toxic doses, especially in neonates. Other studies have shown also that although pediatric-attenuated shocks are ideal, including in the treatment of infants, nonattenuated adult-dosage biphasic shocks still are highly effective and relatively safe, even for a newborn infant.

Because of the prohibition of energy doses higher than the 4 J/kg dose in the past, many children have not received timely shocks for VF, even when shocks were advised by AEDs. It is unknown how many children may have died awaiting defibrillation while receiving lower, recommended energy doses. All available data indicate that traditional and even very high defibrillation doses from AEDs designed for adults are effective and safe in this population and fall below the TD$_{50}$ and LD$_{50}$ levels.

**INTERNATIONAL AND NATIONAL GUIDELINES**

Several international and national guidelines have included evaluations of the literature on AED use on children as the basis on which recommendations have been formed. A recent international review of resuscitation science and a subsequent consensus on science and treatment recommendations referenced 8 levels of evidence (LOE) to provide the basis of the strength of evidence behind the conclusions made the following statement:

Many but not all AED algorithms have been shown to be sensitive and specific for recognizing shockable arrhythmias in children. A standard AED (“adult” AED with adult pad-cable system) can be used for children older than about 8 years of age and weighing more than about 25 kg. Many manufacturers now provide a method for attenuating the energy delivered to make the AED suitable for smaller children (eg, use of a pad-cable system or an AED with a key or switch to select a smaller dose).

The consensus-on-science statement added (LOE represents levels of evidence of the studies used as defined by the AHA):

The ideal energy dose for safe and effective defibrillation for children is unknown. Extrapolation from adult data and pediatric animal studies suggests that biphasic shocks are at least as effective as monophasic shocks and produce less postshock myocardial dysfunction. One LOE 5 and one LOE 6 study show that an initial monophasic or biphasic shock dose of 2 J/kg generally
terminates pediatric VF. Two pediatric case series (LOE 5) report that doses 4 J/kg (up to 9 J/kg) have effectively defibrillated children under 12 years of age, with negligible adverse effects. In 5 animal studies (LOE 6) large (per kilogram) energy doses caused less myocardial damage in young hearts than in adult hearts. In 3 animal studies (LOE 6) and 1 pediatric case series (LOE 5) a 50 J biphasic dose delivered through a pediatric pad/cable system terminated VF and resulted in survival. One piglet (13–26 kg) study (LOE 6) showed that pediatric biphasic AED shocks (50/75/86 J) terminated VF and caused less myocardial injury and better outcome than adult AED biphasic shocks (200/300/360 J).

As a result of this scientific review, the following treatment recommendation was made:

The treatment of choice for pediatric VF/pulseless VT is prompt defibrillation, although the optimum dose is unknown. For automated defibrillation, we recommend an initial pediatric attenuated dose for children 1 to 8 years of age and up to about 25 kg (55 pounds) and 127 cm (50 inches) in length. There is insufficient information to recommend for or against the use of an AED in infants <1 year of age. A variable dose manual defibrillator or an AED able to recognize pediatric shockable rhythms and equipped with dose attenuation are preferred; if such a defibrillator is not available, a standard AED with standard electrode pads may be used. A standard AED (without a dose attenuator) should be used for children > 25 kg (about 8 years of age) and older adolescent and adult victims.

On the basis of the aforementioned consensus-on-science and treatment recommendations and after an evaluation by its Pediatric Advanced Life Support Committee, the American Heart Association made the following recommendations:

Many AEDs have high specificity in recognizing pediatric shockable rhythms, and some are equipped to decrease the delivered energy to make it suitable for children 1 to 8 years of age. Since the publication of the ECC [Emergency Cardiovascular Care] Guidelines 2000, data has shown that AEDs can be safely and effectively used in children 1 to 8 years of age. However, there is insufficient data to make a recommendation for or against using an AED in infants < 1 year of age. In systems and institutions that care for children and have an AED program, it is recommended that the AED have both a high specificity in recognizing pediatric shockable rhythms and a pediatric dose-attenuating system to reduce the dose delivered by the device. In an emergency, if an AED with a pediatric attenuating system is not available, use a standard AED. Turn the AED on, follow the AED prompts, and resume chest compressions immediately after the shock. Minimize interruptions in chest compressions.

Last, the National Association of EMS Physicians also reviewed the evidence that supports the use of AEDs on children and made the following recommendations:

Strategies for treatment of pediatric arrest should focus on shortening the intervals from collapse to recognition of ventricular fibrillation and to defibrillation. Data on the correct energy for defibrillation of children is limited. Animal studies suggest the immature heart is less susceptible to energy related damage than the adult heart and that there is a wide therapeutic range of defibrillation energy dose. Although using a fixed energy AED in some children may have the potential for harm, not treating ventricular fibrillation has the potential for even greater harm, death of the child. As such, defibrillation should not be withheld based on weight and size criteria alone. Systems should attempt to provide defibrillation to children suffering ventricular fibrillation in the timeliest fashion possible. Strategies may include: manual defibrillation, AEDs designed for defibrillation of young children, and standard AEDs used in children with appropriate protocols and medical oversight.

AED DESIGN CONSIDERATIONS

One factor that has delayed professional organization recommendations and protocol development that advocate AED use for young children and infants has been the need to simplify resuscitation training of the lay public. This would necessitate a device designed to detect and analyze the rhythms in children accurately while not requiring dose-energy adjustments on the basis of the weight of the patient or allowing for energy-dose adjustment for children to apply across large age ranges with a minor and easy-to-teach modification.

Some have postulated that any age-specific cutoff, although arbitrary but allowing simplicity, outweighs the need for strict adherence to dosing recommendations. Studies have shown that keeping resuscitation instructions simple provides for improved skill mastery and retention. The more complex the teaching sequence or message, the less likely it is that the rescuer will remember what to do and successfully complete the task. Because lay responders or first responders with limited training in arrhythmia recognition use AEDs, it is important to keep the user interface simple. This principle must be applied to pediatric AEDs as well as adult AEDs. Adding complexity to the AED user interface to accommodate pediatric defibrillation could result in opportunities for error in both adult and pediatric defibrillation. Therefore, any enhancement to allow treatment of pediatric patients with the AED must be as simple as possible while not compromising adult care. If this goal cannot be achieved in certain situations, it may be more appropriate to allow existing adult AEDs to be used on children.

IMPLEMENTATION OF A LAY RESCUE AED PROGRAM IN SCHOOLS WITH A DOCUMENTED NEED

The implementation of an AED program in schools has become a point of major discussion in recent years. Despite this growing interest in the use of AEDs on children, as recommended in a recent American Academy of Pediatrics—endorsed policy statement from the American Heart Association, lay rescuer and emer-
gency preparedness programs should be directed at the complete planning and response to cardiac arrest and other life-threatening conditions rather than focused on a single piece of equipment. The policy-statement recommendations were to establish a comprehensive emergency response plan, which would include the following key elements:

1. Effective and efficient communication throughout the school campus: Establish a rapid communication system that links all parts of the school campus, including outdoor facilities and practice fields, to the EMS system. Establish protocols to clarify when the EMS system and other emergency contact people should be called. Determine the time required for EMS response to any location on campus and establish a method to efficiently direct EMS personnel to any location on campus. Create a list of important contact people and telephone numbers with a protocol to indicate when each person should be called. Include names of experts to help with postevent support.

2. Coordinated and practiced response plan: Develop a response plan for all medical emergencies in consultation with the school nurse, the school or school athletic team physicians, athletic trainers, and the local EMS agency as appropriate. EMS and emergency dispatchers (911 centers) should be made aware of the type of rescue equipment available at the school and its location. Practice the response sequence at the beginning of each school year and periodically throughout the year and evaluate and modify it as needed.

3. Risk reduction: Prevent injuries through safety precautions in classrooms and on the playground. Identify students, faculty, and staff with medical conditions that place them at risk for development of life-threatening conditions and train and equip personnel to provide the appropriate response for those conditions.

4. Training and equipment for first aid and CPR: Ensure that a sufficient number of teachers are trained as CPR and first aid instructors. Train school staff and graduating high school students for CPR. Teachers and staff trained for first aid should, at a minimum, be equipped and able to give first aid for the following life-threatening emergencies until EMS rescuers arrive:
   a. severe breathing problems including asthma, choking, and anaphylaxis (severe allergic reaction);
   b. chest pain and heart attack;
   c. diabetes and low blood sugar;
   d. stroke;
   e. seizure;
   f. shock;
   g. bleeding;
   h. head and spine injury;
   i. broken bones;
   j. burns;
   k. sudden cardiac arrest;
   l. temperature-related emergencies (heatstroke and hypothermia); and
   m. poisoning.

5. Implementation of a lay rescuer AED program in schools with an established need: If the school determines that a lay rescuer AED program is needed, school administrators and medical personnel should include the AED program in the school medical emergency response plan and practice and evaluate response to sudden cardiac arrest with the AED. EMS and 911 centers should be notified of the specific type of AED and the exact location of the AED on the school grounds. Rescuers who are unfamiliar with the school can call 911 and receive instructions from 911 dispatchers to find and use the AED. AED programs should have the following elements:
   a. medical/health care provider oversight;
   b. appropriate training of anticipated rescuers in CPR and use of the AED;
   c. coordination with the EMS system;
   d. appropriate device maintenance; and
   e. an ongoing quality improvement program.

As was discussed in the aforementioned policy statement and has been shown through several studies, the AED may be part of a school emergency response program but should be implemented only as part of a comprehensive program as described, with appropriate oversight, and based on determination of need. To determine the need for an AED program at any location, the policy statement recommended consideration of lay rescuer AED program implementation in locations with at least 1 of the following characteristics:

1. The frequency of cardiac arrest events is such that there is a reasonable probability of AED use within 5 years of rescuer training and AED placement. This probability is calculated on the basis of 1 cardiac arrest known to have occurred at the site within the last 5 years, or the probability can be estimated on the basis of population demographics.

2. There are children attending the school or adults working at the school who are thought to be at high risk for sudden cardiac arrest (eg, children with conditions such as congenital heart disease and a history
of abnormal heart rhythms, children with long QT syndrome, children with cardiomyopathy, adults or children who have had heart transplants, adults with a history of heart disease).

3. An EMS call-to-shock interval of less than 5 minutes cannot be reliably achieved with conventional EMS services and a collapse-to-shock interval of less than 5 minutes can be reliably achieved (in >90% of cases) by training and equipping laypersons to function as first responders by recognizing cardiac arrest, telephoning 911 (or other appropriate emergency response number), starting CPR, and attaching/operating an AED.

When funds are limited but there remains a desire to establish some AED school programs, priority should be given to establishing programs in large schools, schools used for community gatherings, schools at the greatest distance from EMS response, and schools that are attended by the largest number of adolescents and adults (eg, high schools and trade schools).

The 5 key components of an AED program are:

1. medical/health care provider oversight;
2. appropriate training of anticipated rescuers in CPR and use of the AED;
3. coordination with the EMS system;
4. appropriate device maintenance; and
5. an ongoing quality improvement program to monitor training and evaluate response with each use of the device.

If an AED program is established at the school, the AED should be placed in a central location that is accessible at all times and ideally no more than a 1- to 11/2-minute walk from any location. The device should be secure and located near a telephone (eg, near the school office, library, or gymnasium) so that a rescuer can activate the EMS system and get the AED at the same time. The EMS system should be notified of the establishment of the AED program and the emergency medical dispatcher should know the specific type of AED at the school and where it is located. Several staff members should be trained in both CPR and use of the AED.

Recent federal legislation provides guidance for AED programs in schools. HR 389/Pub L No. 108–41 enabled the development of a national resource center to provide schools with information and technical guidance to set up AED programs, giving schools access to the appropriate training, fund-raising techniques, and other logistics required to make such programs successful. The national resource center is modeled after Project ADAM, a joint venture between the Children’s Hospital of Wisconsin and David Ellis, a friend of the project’s namesake, Adam Lemel, who collapsed and died during a high school basketball game. Senate Bill 231 is a companion measure.

**CONCLUSIONS**

Although the incidence of VF in the pediatric population is low, there is a need for developing strategies to provide early defibrillation to patients younger than 8 years. This may include the need for an AED suitable for use on pediatric patients. Because of the limited nature of effective energy-dose data, EMS systems, medical directors, and pediatric researchers should make efforts to gather information regarding pediatric uses of their devices and report it by using the pediatric Utstein style, which represents an internationally accepted standard method of collecting and reporting respiratory and cardiac arrest and resuscitation data. In addition, because the literature suggests that some emergency responders may fear using AEDs on children, EMS and physician leaders should work with professional organizations, community organizations, and researchers to educate both first responders and community members regarding the benefits of early pediatric defibrillation and the use of available varieties of AEDs.

Current pediatric protocols and guidelines recommend energy doses of 2 to 4 J/kg for defibrillation of children. These recommendations evolved from limited data that focused on the likelihood of effectiveness without consideration for therapeutic window determination or analysis of potential toxicity. In addition, these dose recommendations were made on the basis of data from defibrillation with monophasic damped sine waveforms and without impedance compensation, which makes their extrapolation to current monophasic and biphasic technology unreliable. The existing data on the relationship between size or weight and impedance in children are poor, which indicates that weight-based dosing may be of limited value in pediatric defibrillation.

The most important safety feature of an AED is specificity. As long as there is a very high level of assurance that shocks will be advised only for appropriate rhythms in the pediatric population, then the risk of myocardial damage from defibrillation likely is significantly less than the risk of not delivering a shock (probable death). However, the potentially toxic effects of delivering too much energy must be minimized whenever possible. Data extrapolated from animal models support the use of adult-energy AEDs even in smaller pediatric patients.

In the future, technology probably will provide for the development of a “1-size-fits-all” version of an AED. While that technology is being developed, an AED in use today ideally should have both a high specificity in recognizing pediatric shockable rhythms and a pediatric dose-attenuating system to reduce the dose delivered for
children younger than 8 years, including infants. Several AEDs that are currently on the market have such pediatric-attenuating devices that have been approved by the Food and Drug Administration for both children and infants. These same devices have the specificity needed to recognize shockable rhythms in children and infants. As such, these systems are the preferred treatment for both children and infants. However, if an AED with a pediatric-attenuating system is not available, the responder should use a standard AED rather than delay the delivery of a potentially life-saving intervention.

The message for the public and EMS systems is that the existence of VF in children and infants needs to be recognized, and effective methods to treat VF need to be used as early as possible to improve the chance of survival for children and infants after sudden cardiac arrest. In addition, this possibly life-saving therapy should not be withheld purely on the basis of absolute weight and size issues. In locations where AEDs are currently deployed, the acquisition and deployment of models that have the ability to provide an attenuated adult dosage to treat children and infants should be encouraged. In the absence of these attenuated-dose devices, rescuers need to be aware that they should still provide care to infants and children with a nonattenuated adult-dosage device, because the potential for benefit far outweighs the risk. The key is to pursue a long-term goal of providing devices that will allow rapid defibrillation for adult and pediatric patients. This can be accomplished through deploying devices that treat children without compromising adult care, having approaches that minimize device training issues, and optimizing the use of limited financial and personnel resources.

COMMITTEE ON PEDIATRIC EMERGENCY MEDICINE, 2005–2006
Steven E. Krug, MD, Chairperson
Thomas Bojko, MD, MS
Margaret A. Dolan, MD
Karen S. Frush, MD
Patricia J. O’Malley, MD
Robert E. Sapien, MD
Kathy N. Shaw, MD, MSCE
Joan Shook, MD, MBA
Paul E. Sirbaugh, DO
Loren G. Yamamoto, MD, MPH, MBA

LIAISONS
Jane Ball, RN, DrPH
- EMSC National Resource Center
Kathleen Brown, MD
- National Association of EMS Physicians
Kim Bullock, MD
- American Academy of Family Physicians
Dan Kavanaugh, MSW
- Maternal and Child Health Bureau

Sharon E. Mace, MD
- American College of Emergency Physicians
Susan Eads Role, JD, MLSLS
- EMSC National Resource Center
David W. Tuggle, MD
- American College of Surgeons
Tina Turgel, BSN, RN,C
- Maternal and Child Health Bureau

CONTRIBUTORS
David Markenson, MD
Steve Neish, MD
Lee Pyles, MD

STAFF
Susan Tellez

SECTIONS ON CARDIOLOGY AND CARDIAC SURGERY, 2006–2007
Robert H. Beekman III, MD, Chairperson
Peter B. Manning, MD
Seema Mital, MD
William R. Morrow, MD
Frank M. Galioto, Jr, MD
Thomas K. Jones, MD

LIAISONS
Gerard R. Martin, MD
- American College of Cardiology
Reginald L. Washington, MD
- NCE Planning Group

STAFF
Lynn Colegrove, MBA

REFERENCES


