ARTICLES

**Parent Use of Touchscreen Computer Kiosks for Child Health Promotion in Community Settings**
Darcy A. Thompson, Paula Lozano, and Dimitri A. Christakis

**Screening for Depression in an Urban Pediatric Primary Care Clinic**
Howard Dubowitz, Susan Feigelman, Wendy Lane, Leslie Prescott, Kenneth Blackman, Lawrie Grube, Walter Meyer, and J. Kathleen Tracy

**Factors Associated With Identification and Management of Maternal Depression by Pediatricians**

**The Influence of the Timing of Cord Clamping on Postnatal Cerebral Oxygenation in Preterm Neonates: A Randomized, Controlled Trial**
Oskar Baenziger, Florian Stolkin, Mathias Keel, Kurt von Siebenthal, Jean-Claude Fauchere, Seema Das Kundu, Vera Dietz, Hans-Ulrich Bucher, and Martin Wolf

**A Randomized, Controlled Trial of Acetaminophen, Ibuprofen, and Codeine for Acute Pain Relief in Children With Musculoskeletal Trauma**
Eric Clark, Amy C. Plint, Rhonda Correll, Isabelle Gaboury, and Brett Passi

**Is Nephrocalcinosis in Preterm Neonates Harmful for Long-term Blood Pressure and Renal Function?**
Joana E. Kist-van Holthe, Paul H.T. van Zwieten, Eveline A. Schell-Feith, Harmien M. Zonderland, Herma C. Holscher, Ron Wolterbeek, Sylvia Veen, Marijke Frolich, and Bert J. van der Heijden

**Steroid Use Among Adolescents: Longitudinal Findings From Project EAT**
Patricia vandenBerg, Dianne Neumark-Sztainer, Guy Cafri, and Melanie Wall

**Patient and Hospital Correlates of Clinical Outcomes and Resource Utilization in Severe Pediatric Sepsis**
Folafoluwa O. Odetola, Achamyelah Gebremariam, and Gary L. Freed

**Risk of Recurrent Childhood Arterial Ischemic Stroke in a Population-Based Cohort: The Importance of Cerebrovascular Imaging**
Heather J. Fullerton, Yvonne W. Wu, Stephen Sidney, and S. Claiborne Johnston

**Auditory Neural Maturation After Exposure to Multiple Courses of Antenatal Betamethasone in Premature Infants as Evaluated by Auditory Brainstem Response**
Sanjiv B. Amin and Ronnie Guillet

**Stress Predicts Brain Changes in Children: A Pilot Longitudinal Study on Youth Stress, Posttraumatic Stress Disorder, and the Hippocampus**
Victor G. Carrion, Carl F. Weems, and Allan L. Reiss
Twelve-Month Effectiveness of a Parent-led, Family-Focused Weight-Management Program for Prepubertal Children: A Randomized, Controlled Trial
Rebecca K. Golley, Anthea M. Magarey, Louise A. Baur, Katharine S. Steinbeck, and Lynne A. Daniels

Work-Related Hazards and Workplace Safety of US Adolescents Employed in the Retail and Service Sectors
Carol W. Runyan, Michael Schulman, Janet Dal Santo, J. Michael Bowling, Robert Agans, and Myduc Ta

Infant Care Patterns at Epidemiologic Study of Cystic Fibrosis Sites That Achieve Superior Childhood Lung Function
Raj Padman, Susanna A. McColley, Dave P. Miller, Michael W. Konstan, Wayne J. Morgan, Michael S. Schechter, Clement L. Ren, Jeffrey S. Wagener for the Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis

Inaccuracy in Pediatric Outpatient Blood Pressure Measurement
Amber Podoll, Michelle Grenier, Beth Croix, and Daniel I. Feig

Preventive Dental Care for Children in the United States: A National Perspective
Charlotte W. Lewis, Brian D. Johnston, Kristi A. Linsenmeyer, Alexis Williams, and Wendy Mouradian

Postterm Delivery and Risk for Epilepsy in Childhood
Vera Ehrenstein, Lars Pedersen, Vibeke Holsteen, Helle Larsen, Kenneth J. Rothman, and Henrik T. Sørensen

Comparison of Current Health, Functional Limitations, and Health Care Use of Young Adults Who Were Born With Extremely Low Birth Weight and Normal Birth Weight
Saroj Saigal, Barbara Stoskopf, Michael Boyle, Nigel Paneth, Janet Pinelli, David Streiner, and John Goddeeris

Measles-Mumps-Rubella and Varicella Vaccine Responses in Extremely Preterm Infants
Carl T. D'Angio, Paulina A. Boohene, Anne Mowrer, Susette Audet, Marilyn A. Menegus, D. Scott Schmid, and Judy A. Beeler

Influenza Vaccine Coverage and Missed Opportunities Among Inner-city Children Aged 6 to 23 Months: 2000-2005
Jennifer R. Verani, Matilde Irigoyen, Shaofu Chen, and Frank Chimkin

Vaccine Effectiveness Against Medically Attended, Laboratory-Confirmed Influenza Among Children Aged 6 to 59 Months, 2003–2004
Carrie M. Shuler, Martha Iwamoto, Carolyn Buxton Bridges, Mona Marin, Ruth Neeman, Paul Gargiullo, Terrace A. Yoder, Harry L. Keyserling, and Pauline D. Terebuh

Toward Creating Family-Friendly Work Environments in Pediatrics: Baseline Data From Pediatric Department Chairs and Pediatric Program Directors
Heather A. McPhillips, Ann E. Burke, Kate Sheppard, Adam Pallant, F. Bruder Stapleton, and Bonita Stanton

Incidence, Prognosis, and Risk Factors for Fatigue and Chronic Fatigue Syndrome in Adolescents: A Prospective Community Study
Katharine A. Rimes, Robert Goodman, Matthew Hotopf, Simon Wessely, Howard Meltzer, and Trudie Chalder

Early Childhood Gender Differences in Anterior and Posterior Cerebral Blood Flow Velocity and Autoregulation
Nuj Tontisirin, Saipin L. Muangman, Pilar Suz, Catherine Pihoker, Dana Fisk, Anne Moore, Arthur M. Lam, and Monica S. Vavilala
Neurodevelopmental Consequences of Early Traumatic Brain Injury in 3-Year-Old Children
Heather T. Keenan, Stephen R. Hooper, Crista E. Wetherington, Maryalice Nocera, and Desmond K. Runyan

Weight Status in Young Girls and the Onset of Puberty
Joyce M. Lee, Danielle Appugliese, Niko Kaciroti, Robert F. Corwyn, Robert H. Bradley, and Julie C. Lumeng

Pediatric and Emergency Medicine Residents' Attitudes and Practices for Analgesia and Sedation During Lumbar Puncture in Pediatric Patients
Vicky R. Breakey, Jonathan Pirie, and Ran D. Goldman

How Children With Special Needs Travel With Their Parents: Observed Versus Reported Use of Vehicle Restraints
Taube Korn, Michal Katz-Leurer, Shirley Meyer, and Rosa Gofin

Admission Temperature of Low Birth Weight Infants: Predictors and Associated Morbidities
Abbot R. Laptook, Walid Salhab, Brinda Bhaskar and the Neonatal Research Network

Lead Exposure, IQ, and Behavior in Urban 5- to 7-Year-Olds: Does Lead Affect Behavior Only by Lowering IQ?
Aimin Chen, Bo Cai, Kim N. Dietrich, Jerilynn Radcliffe, and Walter J. Rogan

Mortality of Late-Preterm (Near-Term) Newborns in Utah
Paul C. Young, Tiffany S. Glasgow, Xi Li, Ginger Guest-Warnick, and Gregory Stoddard

Transmission of West Nile Virus Through Human Breast Milk Seems to Be Rare
Alison F. Hinckley, Daniel R. O'Leary, and Edward B. Hayes

Influence of Multiple Risk Behaviors on Physical Activity–Related Injuries in Adolescents
Ian Janssen, Suzanne Dostaler, William F. Boyce, and William Pickett

Neurodevelopmental Functioning in HIV-Infected Infants and Young Children Before and After the Introduction of Protease Inhibitor–Based Highly Active Antiretroviral Therapy
Jane C. Lindsey, Kathleen M. Malee, Pim Brouwers, Michael D. Hughes for the PACTG 219C Study Team

Infectious Disease Morbidity Among Young HIV-1–Exposed But Uninfected Infants in Latin American and Caribbean Countries: The National Institute of Child Health and Human Development International Site Development Initiative Perinatal Study
Marisa M. Mussi-Pinhata, Laura Freimanis, Aparecida Y. Yamamoto, James Korelitz, Jorge A. Pinto, Maria L. S. Cruz, Marcelo H. Losso, Jennifer S. Read for the National Institute of Child Health and Human Development International Site Development Initiative Perinatal Study Group

Once-Daily Highly Active Antiretroviral Therapy for HIV-Infected Children: Safety and Efficacy of an Efavirenz-Containing Regimen
Henriëtte J. Scherpbier, Vincent Bekker, Daşja Pajkrt, Suzanne Jurriaans, Joep M. A. Lange, and Taco W. Kuijpers

Early-Life Risk Factors for Occurrence of Atopic Dermatitis During the First Year
Mikio Sugiyama, Hirokazu Arakawa, Kiyoshi Ozawa, Takahisa Mizuno, Hiroyuki Mochizuki, Kenichi Tokuyama, and Akihiro Morikawa
Bacterial Imprinting of the Neonatal Immune System: Lessons From Maternal Cells?
Pablo F. Perez, Joël Doré, Marion Leclerc, Florence Levenez, Jalil Benyacoub, Patrick Serrant, Iris Segura-Roggiero, Eduardo J. Schiffrin, and Anne Donnet-Hughes

Evaluation of Psychopathological Conditions in Children With Heavy Prenatal Alcohol Exposure
Susanna L. Fryer, Christie L. McGee, Georg E. Matt, Edward P. Riley, and Sarah N. Mattson

Di-(2-ethylhexyl)phthalate and Deep Venous Thrombosis in Children: A Clinical and Experimental Analysis

STATE-OF-THE-ART REVIEW ARTICLE

The Ketogenic Diet: One Decade Later
John M. Freeman, Eric H. Kossoff, and Adam L. Hartman

REVIEW ARTICLES

Has Blood Pressure Increased in Children in Response to the Obesity Epidemic?
Arnaud Chiolero, Pascal Bovet, Gilles Paradis, and Fred Paccaud

Childhood Helicobacter pylori Infection and Growth Impairment in Developing Countries: A Vicious Cycle?
Henry J. Windle, Dermot Kelleher, and Jean E. Crabtree

SPECIAL ARTICLES

The Late Effects of Childhood Cancer Therapy
Joseph D. Dickerman

Internet-Based Home Monitoring and Education of Children With Asthma Is Comparable to Ideal Office-Based Care: Results of a 1-Year Asthma In-Home Monitoring Trial
Debora S. Chan, Charles W. Callahan, Virginia B. Hatch-Pigott, Annette Lawless, H. Lorraine Proffitt, Nola E. Manning, Mary Schweikert, and Francis J. Malone

Predictors of Pain and/or Fever at 3 to 7 Days for Children With Acute Otitis Media Not Treated Initially With Antibiotics: A Meta-analysis of Individual Patient Data
Maroëska M. Rovers, Paul Glasziou, Cees L. Appelman, Peter Burke, David P. McCormick, Roger A. Damoiseaux, Paul Little, Nicole Le Saux, and Arno W. Hoes
COMMENTARIES

Pain Assessment in Preterm Neonates
K. J. S. Anand

P Less Than .05: What Does It Really Mean?
Zeev N. Kain and Jill MacLaren
Pediatrics 2007; 119: 608-610.

Oxygen Therapy for Bronchiolitis
Joel L. Bass and David Gozal
Pediatrics 2007; 119: 611.

Pediatric Palliative Care: A Novel Approach to Children With Sickle Cell Disease
Brenda C. McClain and Zeev N. Kain

The Use of Amplitude-Integrated Electroencephalography: Beware of Its Unintended Consequences
John M. Freeman

FROM THE AMERICAN ACADEMY OF PEDIATRICS

Cardiovascular Risk Reduction in High-Risk Pediatric Populations

High-Deductible Health Plans and the New Risks of Consumer-Driven Health Insurance Products
Committee on Child Health Financing

Testing for Drugs of Abuse in Children and Adolescents: Addendum—Testing in Schools and at Home
Committee on Substance Abuse and Council on School Health
Pediatrics 2007; 119: 627-630.

Special Requirements of Electronic Health Record Systems in Pediatrics
S. Andrew Spooner and the Council on Clinical Information Technology

Self-injectable Epinephrine for First-Aid Management of Anaphylaxis
Scott H. Sicherer, F. Estelle R. Simons and the Section on Allergy and Immunology

LETTERS TO THE EDITOR

Pulmonary Manifestations in ΔF508/R117H
Clement L. Ren

Pulmonary Manifestations in ΔF508/R117H: In Reply
Brian P. O'Sullivan, Anne Marie Comeau, and Richard Parad

Unplanned Transport Events and Severity of Illness: Are We Conveying the Whole Picture?
Bradley A. Kuch, Ricardo Munoz, Richard A. Orr, and R. Scott Watson
Unplanned Transport Events and Severity of Illness: Are We Conveying the Whole Picture?: In Reply
Scott B. Yeager and Peter C. Laussen

Industry-Sponsored "Expert Committee Recommendations for Acne Management"
Promote Expensive Drugs on the Basis of Weak Evidence
Thomas B. Newman

Industry-Sponsored "Expert Committee Recommendations for Acne Management"
Promote Expensive Drugs on the Basis of Weak Evidence: In Reply
Andrea L. Zaenglein and Diane M. Thiboutot

Preparing for Asthma-Related Emergencies in Schools
Robert E. Sapien

Preparing for Asthma-Related Emergencies in Schools: In Reply
Marianne Hillemeier, Yu Bai, and Maryellen E. Gusic

Screening for Developmental Dysplasia of the Hip
Perry L. Schoenecker and John M. Flynn

Screening for Developmental Dysplasia of the Hip: In Reply
Ned Calonge and Diana Petti

Payment for Telephone Care: From Policy to Practice
Randall S. Sterkel

Payment for Telephone Care: From Policy to Practice: In Reply
Sanford M. Melzer

Improving Developmental Screening: Combining Parent and Pediatrician Opinions With Standardized Questionnaires
Paul H. Lipkin and Hilary Gwynn

Improving Developmental Screening: Combining Parent and Pediatrician Opinions With Standardized Questionnaires: In Reply
Michael I. Shevell

Support for Addressing the Effects of Public Policy on Same-Gender Couples and Their Children
Lee Ann Savio Beers, Nathaniel Beers, and Benjamin Gitterman

Thirteen-Year Spontaneous Evolution of Helicobacter pylori Gastritis Acquired During Early Childhood
Léonie Martigne, Laurent Michaud, Frédéric Gottrand, Françoise Boman, and Marie-Odile Husson

ERRATA

ERRATA
Recurrent Expressive Aphasia as a Presentation of Cat-Scratch Encephalopathy
James W. Fox, Joanna K. Studley, and Daniel M. Cohen

Marked Hyperbilirubinemia Associated With the Heme Oxygenase-1 Gene Promoter Microsatellite Polymorphism in a Boy With Autoimmune Hemolytic Anemia
Stephan Immenschuh, Ying Shan, Hartmut Kroll, Sentot Santoso, Wilhelm Wössmann, Gregor Bein, and Herbert L. Bonkowsky

Radiologic and Neurophysiologic Aspects of Stroke-like Episodes in Children With Congenital Disorder of Glycosylation Type Ia
Argirios Dinopoulos, Ismail Mohamed, Blaise Jones, Sanjai Rao, David Franz, and Ton deGrauw

Citrin Deficiency: A Novel Cause of Failure to Thrive That Responds to a High-Protein, Low-Carbohydrate Diet
David Dimmock, Keiko Kobayashi, Mikio Iijima, Ayako Tabata, Lee-Jun Wong, Takeyori Saheki, Brendan Lee, and Fernando Scaglia

Sinus Bradycardia After Intravenous Pulse Methylprednisolone
Jonathan D. Akikusa, Brian M. Feldman, Gil J. Gross, Earl D. Silverman, and Rayfel Schneider

Scurvy and Rickets Masked by Chronic Neurologic Illness: Revisiting "Psychologic Malnutrition"
James McCallum Noble, Arthur Mandel, and Marc C. Patterson
ABSTRACT

OBJECTIVES. The goals were to evaluate the use of touchscreen computer kiosks, containing only child health–promoting information, in urban, low-income, community settings and to characterize the users of these kiosks.

METHODS. Three user-driven touchscreen computer kiosks were placed in low-income urban locations in Seattle, Washington, from March 2005 to October 2005. The locations included a public library, a Department of Motor Vehicles office, and a McDonald’s restaurant. Users selected age-appropriate modules with prevention information and screening tools. Users entered the age of the child and were presented with age-appropriate modules. On exiting, users were asked to rate their experience and to provide basic demographic data.

RESULTS. In total, there were 1846 kiosk sessions. Almost one half occurred at McDonald’s. Seventy-eight percent of users identified themselves as first-time users. Users sought information for children of all ages. Sixty-one percent of first-time users explored 1 module. First-time users were most interested in television/media use (16%), smoke exposure (14%), attention-deficit/hyperactivity disorder screening (12%), and asthma assessment (11%). At-risk children were identified in 52% of sessions. Eighty-seven percent of first-time users who completed the asthma assessment had children whose asthma was uncontrolled. Twenty-eight percent of users responded to ≥1 question on the exit survey. Of those, 48% had less than a high school education, and 26% had never used the Internet. Approximately one half found the kiosk easy to use (57%) and the information easy to understand (55%); 66% said there was at least some new information. Fifty-five percent planned to try some of the things they had learned, and 49% intended to talk to their child’s doctor about what they had learned.

CONCLUSIONS. User-driven computer kiosks were used in community settings to obtain child health information. Users found the kiosks easy to use. Additional study on improving use and understanding the impact is needed.
Given the limited success in promoting prevention via pediatricians’ offices, alternative methods for delivering child health promotion information are needed. Although recommendations regarding age-appropriate topics to be covered at well-child visits are plentiful, parents report that many of the anticipatory guidance topics are not covered in their children’s well-child visits and that, after their visits, they can still use more information. In addition, many parents have little interaction with health care providers. Children >2 years of age may visit a physician only once per year. Many families face access issues that make it difficult for them to keep up with well-child visits. Therefore, alternative methods for delivery of this information need to be explored.

Technological advances in the past 2 decades have led to a focus on the use of new technologies to assist in the delivery of health promotion messages. One technology now being used is the touchscreen computer kiosk containing health information. Many studies have evaluated the placement of such kiosks in physician office waiting rooms and medical facilities. Very few studies have focused on the pediatric population. Two studies found that kiosks containing child health promotion information had positive effects on parental knowledge of prevention topics. It is notable, however, that the content of the kiosks in both of those studies was limited; one focused on household safety topics and the other on topics relevant to infants <6 months of age.

Because health promotion information must be provided outside medical facilities (for a variety of reasons), there has been a push to place health kiosks in the community. Two of the largest efforts are the Michigan Interactive Health Kiosk project and the United Kingdom-based In Touch With Health project, both of which placed >100 kiosks in community locations. The Michigan Interactive Health Kiosk project presented mainly adult-focused information, with very limited child-focused information. The In Touch With Health project covered a broader range of pediatric issues and found that pediatric topics were among the top 20 medical topics viewed. Evaluation of the general use of these community-based kiosks showed that the kiosks were well used, including use by low-income and diverse populations.

To our knowledge, however, there has not been an evaluation of the use of community-based computer kiosks that contain only child health-promoting information. In addition, there has not been a detailed description of the users who seek child health-promoting information from community-based kiosks. The objectives of this study were therefore to evaluate the use of touchscreen computer kiosks containing only child health-promoting information, in urban, low-income, community settings, and to characterize the users of such kiosks.

**METHODS**

**Setting**

This was a descriptive study of community-based kiosk use. The study protocol was approved by the University of Washington institutional review board.

Between March 2005 and October 2005, we placed 3 touchscreen computer kiosks at sites in low-income urban neighborhoods in Seattle, Washington. Sites included a McDonald’s restaurant, a Department of Motor Vehicles (DMV) office, and a public library. Data from the corresponding US Census tracts for these locations showed that 20% to 25% of the population in these communities is <14 years of age, 26% to 32% of parents are single, and 49% to 65% of adults have a high school education or less. The specific location of the kiosk at each site was decided by the site managers, in consultation with the investigators.

**Kiosk Content and Design**

Kiosks contained 17-inch, flat, touchscreen, cathode-ray tube monitors. Each computer kiosk, including computer, touchscreen, and kiosk, cost $2200. To decrease maintenance needs, the kiosks were configured to be rebootable and to reload automatically. We labeled the kiosks as containing child health information, both on the screen and on the sides of the kiosks. Beyond this, there was no advertising for the kiosks. We adapted software used in 2 clinic-based studies for use in this study. All content was written at the eighth-grade level. We used clearly readable font sizes and included pictures in most modules. All information presented was user-driven, interactive, and presented in stages. All screen touches were recorded. To initiate a kiosk session, users first indicated whether they were previous or new users. A kiosk session was defined as an event in which a person logged onto the kiosk by identifying himself or herself as a new or previous user. Both previous and new users indicated whether they were adults or children. Only data entered by self-identified adults were used for this study. Users selected the month and year of birth for the index child and entered it by tapping “okay.” Kiosks then displayed a list of age-appropriate module topics, which users could enter and exit at any point, viewing as many modules as they wanted. A sample module topics page is shown in Fig 1.

There were 14 modules in total. Ten focused on prevention and safety, including television/media, gun injuries, bicycle injuries, car crashes, tobacco smoke exposure, flu shots, sudden infant death syndrome prevention, house fires, Head Start, and scald burn prevention. Three modules were screening tools for the following: developmental delay, tuberculosis, and attention-deficit/hyperactivity disorder (ADHD). The final module was a symptom assessment tool for children with asthma. Users were shown only modules that were age-appropriate for the child of interest. The number of
module topics listed on the topics page ranged from 9 to 12, depending on the age of the child (Table 1).

Content was derived from age-specific recommendations obtained from 5 sources: (1) The United States Preventive Services Task Force Guide to Clinical Preventive Services,21 (2) Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents,4 (3) National Asthma Education and Prevention Program Expert Panel Report 2,22 (4) peer-reviewed systematic reviews of other preventive care interventions, and (5) high-quality randomized, controlled trials. We selected topics for which there were reasonable evidence bases and clear recommenda-
tions. None of the content was specific to the Seattle communities in which this study was performed.

In a subset of 9 modules, users were asked specific questions at the start of the module. An algorithm was applied to identify whether the index child was at risk with respect to the module topic. Users responded to questions to describe the risk level of their child with respect to 5 prevention topics, namely, television/media, tobacco smoke exposure, bicycle injuries, house fire prevention, and sudden infant death syndrome prevention, and all 4 of the screening/assessment modules. For example, in the bicycle injuries module, users were asked

![Figure 1](image-url)

**Sample screen page (module topics page shown to parents of 5-year-old children).** TB indicates tuberculosis; TV, television.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Modules Listed on Topics Page, According to Age of Child</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;6 mo</td>
</tr>
<tr>
<td>ADHD screen</td>
<td></td>
</tr>
<tr>
<td>Asthma assessment</td>
<td>X</td>
</tr>
<tr>
<td>Bicycle injuries</td>
<td>X</td>
</tr>
<tr>
<td>Car crashes</td>
<td>X</td>
</tr>
<tr>
<td>Developmental delay screen</td>
<td>X</td>
</tr>
<tr>
<td>Flu shot</td>
<td>X</td>
</tr>
<tr>
<td>Gun injuries</td>
<td>X</td>
</tr>
<tr>
<td>Head Start</td>
<td>X</td>
</tr>
<tr>
<td>House fire prevention</td>
<td>X</td>
</tr>
<tr>
<td>Scald burns prevention</td>
<td>X</td>
</tr>
<tr>
<td>SIDS prevention</td>
<td>X</td>
</tr>
<tr>
<td>Tobacco smoke exposure</td>
<td>X</td>
</tr>
<tr>
<td>Tuberculosis risk screen</td>
<td>X</td>
</tr>
<tr>
<td>Television/media use</td>
<td>X</td>
</tr>
</tbody>
</table>

SIDS indicates sudden infant death syndrome.

a The bicycle injuries module was not shown for 4-year-old children because of a software programming error discovered at the conclusion of the trial.

b The developmental delay screen was shown for children ≥8 years of age.
3 questions, that is, whether the child rode a bike, (if yes) whether the child had a helmet, and (if yes) whether the child used the helmet when riding a bike. Children reported as not owning or not using a helmet were identified as at-risk children. At-risk criteria for the other 4 prevention topics are shown in Table 2. We used the Vanderbilt screening tool to identify children who warranted additional evaluation for ADHD. The Parents’ Evaluations of Developmental Status (PEDS) screening tool identified children at risk for developmental delays. Screening guidelines from the Centers for Disease Control and Prevention were used to identify children at risk for tuberculosis. We created questions to identify children with inadequately controlled asthma symptoms by using the National Asthma Education and Prevention Program Expert Panel Report 2 guidelines.

After responding to the risk questions, users of these modules were shown messages that depended on their responses. In the majority of these modules, the messages were tailored to the previous responses of the users. For example, in the developmental screening module, if a child was found to be at high risk for a developmental delay, then users were informed of their child’s risk and were encouraged to speak to their child’s physician regarding the findings.

A large button labeled “Exit” was found in the top right corner of all screens and led users to an optional, 9-item, exit survey. The survey asked users to characterize their kiosk experience according to ease of use, newness and helpfulness of the information that was provided, and intent to take action because of the information they viewed (ie, the following yes/no questions: Do you plan to try any of the things that you read about here? Will you talk to your child’s doctor about the things you read today?). The survey also asked the user’s highest grade completed in school, frequency of Internet use, and rating of the child’s health on a 5-point Likert scale (excellent to poor).

Analyses
The total number of adult sessions was summed. We then tabulated frequency distributions for identified categories. Statistical significance was determined by using the $\chi^2$ test unless cell values were $\leq 5$, in which case Fisher’s exact test was used. To reduce the risk of double-counting users, the majority of our analyses were limited to first-time users; we indicate clearly when use by previous users was examined. We also performed multiple Poisson regression analyses with robust error variance to evaluate the relationship between characteristics of users and their ratings of their kiosk session. Poisson regression analysis with robust error variance was used because it is a better way to evaluate relative risk, compared with logistic regression analysis, when the outcome is common. All analyses were performed by using Intercooled Stata 9.1 for Windows (Stata, College Station, TX).

RESULTS
During the 6-month study period, self-identified adults used the touchscreen computer kiosks for 1846 sessions. In 22% of those sessions, users identified themselves as previous users ($n = 399$). Almost one half (47%; $n = 873$) of all kiosk sessions occurred at the McDonald’s restaurant location, with 35% ($n = 642$) occurring at the public library and the remaining 18% ($n = 331$) occurring at the DMV office. Users sought information for children of all ages (Table 3). Almost one half (48%; $n = 117$) of the first-time users who responded to the optional exit questions reported that they had a high school education or less, and one fourth (26%; $n = 67$) reported that they never used the Internet. Forty percent

<table>
<thead>
<tr>
<th>Module</th>
<th>At-Risk Criteria</th>
<th>Viewed Module*</th>
<th>At Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Television/media use</td>
<td>No rules on viewing time</td>
<td>217</td>
<td>123 (57)</td>
</tr>
<tr>
<td>Tobacco smoke exposure</td>
<td>Smoke exposure and no rules to limit exposure</td>
<td>188</td>
<td>73 (39)</td>
</tr>
<tr>
<td>Bicycle injuries</td>
<td>No helmet or no use of helmet</td>
<td>131</td>
<td>53 (40)</td>
</tr>
<tr>
<td>House fire prevention</td>
<td>No smoke detector</td>
<td>69</td>
<td>25 (36)</td>
</tr>
<tr>
<td>SIDS prevention</td>
<td>Stomach or side sleeping</td>
<td>25</td>
<td>14 (56)</td>
</tr>
<tr>
<td>Screening modules</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma assessment</td>
<td>Not well controlled</td>
<td>56</td>
<td>49 (87)</td>
</tr>
<tr>
<td>Tuberculosis risk screen</td>
<td>Special risk for tuberculosis</td>
<td>49</td>
<td>40 (82)</td>
</tr>
<tr>
<td>ADHD screen</td>
<td>Needs additional evaluation</td>
<td>95</td>
<td>38 (40)</td>
</tr>
<tr>
<td>Developmental delay screen</td>
<td>High risk for delay</td>
<td>49</td>
<td>33 (67)</td>
</tr>
<tr>
<td>Total at-risk children identified</td>
<td></td>
<td>712</td>
<td>373 (52)</td>
</tr>
</tbody>
</table>

The 9 modules for which at-risk criteria were defined are shown. SIDS indicates sudden infant death syndrome.

* To be included in the denominator of those who viewed the modules, users needed to complete all of the risk evaluation questions in the named modules.
(n = 103) of first-time users reported their child’s health as fair or poor.

First-time users showed interest in all of the modules presented (Fig 2). The television/media (16%; n = 235), smoke exposure (14%; n = 196), ADHD screening (12%; n = 180), and asthma assessment (11%; n = 155) modules were the most commonly visited modules. For the majority of first-time kiosk sessions, only 1 module was explored (n = 882; 61%); 13% of users viewed 2 modules, and 6% viewed ≥3 modules.

There were 712 first-time user sessions that explored and completed ≥1 of the 9 modules that identified at-risk groups. At-risk children were identified in 373 (52%) of these sessions. Table 2 shows, according to module, the number of sessions in which an at-risk child was identified. Of the 217 first-time users who explored the television/media module, 123 (57%) did not have rules to limit the amount of television their child watched. Of the 188 users who explored the tobacco smoke exposure module, 73 (39%) lived in homes without rules to limit secondhand smoke exposure from smokers in the home. Fifty-three (40%) of the 131 users who viewed the bicycle injuries module sought information for children who ride bicycles and either do not own a helmet or do not wear a helmet when biking. Of the 56 first-time sessions in which the asthma assessment module was completed, the assessment showed that the child’s asthma was not well controlled in 49 (87%) of the sessions. Of the 95 sessions in which the ADHD screening was completed, 38 (40%) identified a child at risk for ADHD.

In the optional exit survey, first-time users rated their kiosk experience. Twenty-eight percent (n = 405) of first-time users responded to ≥1 question on this exit survey, and 11% (n = 163) completed all 9 questions. The majority (57%; n = 155) of those who responded found the kiosk easy to use. Thirty percent found it hard to use, with the rest responding that it was neither hard nor easy. Fifty-five percent (n = 138) rated the information as easy to understand. One fourth (26%; n = 64) found it hard to understand, with the rest stating that it was neither hard nor easy. Ease of use of the kiosk (easy versus not hard, not easy or hard) was associated with Internet use (any versus rarely/never; prevalence ratio: 1.44; 95% confidence interval: 1.09–1.91) but not with above high school education versus less (prevalence ratio: 1.04; 95% confidence interval: 0.81–1.34).

Differences in use according to kiosk location also existed. The kiosk located at the public library had more
previous users (35% of all sessions there) than did the kiosks at the other 2 locations (DMV office: 19%; McDonald’s: 13%; P < .001). First-time users at the DMV office were more likely to explore ≥2 modules (P = .002). However, users at the 3 locations did not vary according to grade level or reported Internet use. In addition, the number of identified at-risk children for first-time users did not vary according to the kiosk location, nor did the selection of modules viewed.

DISCUSSION
Touchscreen computer kiosks located in low-income, urban, community locations were used by users seeking information for children of all ages, on a variety of topics. Notably, users sought information for at-risk children, highlighting their need for information and the potential importance of community-based health kiosks. Overall, users represented a diverse group that included those with low educational levels and a lack of familiarity with the Internet. The majority of users reported a positive experience with their kiosk use, noting that the kiosk was easy to use and contained helpful information.

We placed the kiosks in varied, public, nonmedical locations. The amount of use in the 3 locations varied, with almost one half of the sessions being at the McDonald’s location. To our knowledge, no one has evaluated the public use of a health kiosk placed in a fast food restaurant. Nicholas et al26 noted that a kiosk placed in a supermarket, compared with other medical or nonmedical locations, in the United Kingdom was highly used. The advantages of these locations include high visibility and potentially high volume of use.26 We also found that the number of modules explored varied according to location. First-time users at the DMV office were more likely to explore ≥2 modules. This might be attributable to wait times involved in using services at the DMV office. However, we did not find any difference in the type of user, although this might be attributable simply to the use of only 2 variables (educational level and frequency of Internet use) to characterize viewers. Another study of community-based health kiosks found that users in different locations differed with respect to other characteristics.26

It was notable in our findings that the majority of first-time users responded that they planned to try what they read and almost one half planned to speak with their physician about what they read. This highlights an important strength of community-based health kiosks, which is the ability to provide tailored information to users. Kiosks used in this study were user-driven; users could select the topics that interested them. In addition, information given to users in most modules was customized on the basis of the responses they provided on previous screens. A few studies examined the effect of providing tailored messages versus general health messages and found that customized messages were better accepted by readers and were more likely to have an impact on the readers’ behavior.27–30 The ability for users to select what they read and the customization of some messages on our kiosks may account for our findings and deserve additional exploration.

An additional benefit of community-based child health kiosks is the ability to provide low-cost information to caretakers of at-risk children and adolescents in a manner that is convenient to the user. The majority of costs are start-up costs, including the costs for purchase of the equipment and development of the software. There was no training required for staff members at the community locations. Gould and Anderson31 performed an economic analysis comparing computer kiosks with group sessions with a trainer for nutritional education and found that computer kiosks could reach more individuals at lower costs.

The placement of these kiosks in the community allows users to use the kiosks at a place and time that may be more convenient for them than at a physician’s office. Also, the kiosks provide an additional source of information for those most in need of information, that is, caretakers of at-risk children. More than one half of first-time users who completed risk assessments in the 9 modules with risk questions were identified as having at-risk children. In comparing our findings with national risk levels for some of the specific module findings, we found that users who chose to view each of these modules were more likely to have high-risk children. For example, a recent study found that 20% of parents of infants in Washington state put their child to sleep in a nonsupine position.32 Fifty-six percent of users who viewed the sudden infant death syndrome prevention module in this study reported that they put their child to sleep in a nonsupine position. A national study found that only 3% of households in the United States reported that they did not have a smoke detector.33 Thirty-six percent of kiosk users in our study who chose to view the house fire prevention module reported that they did not have a smoke detector in their home. Therefore, community-based health kiosks allow caretakers of those at most risk to view the information they need.

There are a few important limitations of this study that warrant mention. The most significant of these limitations is that we did not validate the information entered into the computer kiosks. This does not affect our presentation of results related to use, but it may affect the validity of our results related to the risk behaviors of users’ children and our exit survey. Users might have answered module questions inaccurately for their child simply because of curiosity regarding what information might be presented. In addition, some users might have marked responses incorrectly because of difficulty using the kiosk. Thirty percent of users found the kiosks difficult to use. Another limitation is that only 28% of participants entered data in our exit survey.
with 11% answering all of the questions. This hampers our ability to generalize from those users to the larger group of users. Significant differences may exist between those who answered the exit questions and those who did not. In addition, our data do not show whether a child was playing with the kiosk or using it. Users were asked whether they were children or adults; however, children could have been playing with the kiosk and hitting it randomly. We think that this most likely would affect only the overall numbers of sessions we counted for adults. To prevent random hits to the screen affecting our remaining data, users needed to view 4 or 5 different pages and to make 6 to 9 screen touches sequentially to view the module topics page. Any user hitting the screen randomly would have a low likelihood of getting to the modules and answering intramodule questions or viewing the optional exit survey.

An additional limitation is that simply viewing a page or a module does not mean that the viewer is interested in the topic or even reads the information on the screen. Therefore, we do not report pages viewed or time spent viewing the pages. Because viewers had multiple options regarding which modules they viewed, we can assume some interest in most viewers for the modules they viewed. The fact that the majority of viewers viewed only 1 module is consistent with the results of another study. Whether this reflects interest or time constraints is not known.

Despite these limitations, our study shows that community-based, child health kiosks are used by diverse users seeking information for at-risk children. Additional study is needed in a number of areas. Enhancing the ease of use of the kiosk is important, because almost one third of users responded that the kiosk was difficult to use. This is much higher than findings from 2 other community-based kiosk studies, in which >90% of users found the kiosks easy/fairly easy to use. However, both of those studies collected the data on ease of use through in-person interviews, compared with use of the kiosk. Nevertheless, the reasons why one third of the user population had difficulty with the kiosk (ie, software design, unidentified technological problems, factors unique to the users at these locations, or some other reason) need to be explored before widespread distribution. It would also be important to know whether the users who reported that they intended to act on what they learned actually followed through with this. A recent study that was more home Internet based examined similar module content and parents’ implementation of a health promotion change and found that viewing the content did lead to reported behavior change. More attention should be given to those in this study who did not report that they intended to act on this information; understanding why would be important. Similarly, it must be asked whether additional tailoring of the information in the kiosk would have stimulated more parents to report an intention to act. In general, more emphasis needs to be placed on improving child health through effective methods outside medical facilities. Community-based health kiosks with child health information have the potential to have an impact on health care and health outcomes for children.

ACKNOWLEDGMENTS

This study was funded by a royalty research grant from the University of Washington (to Dr Christakis) and by a grant from the Agency for Healthcare Research and Quality (grant 1 R01 HS013302, to Dr Christakis). Support for this study for Dr Thompson was provided by the Robert Wood Johnson Foundation through the Robert Wood Johnson Clinical Scholars Program.

REFERENCES

Screening for Depression in an Urban Pediatric Primary Care Clinic

Howard Dubowitz, MD, MSa, Susan Feigelman, MDa, Wendy Lane, MD, MPHa,b, Leslie Prescott, BAa, Kenneth Blackman, MSa, Lawrie Grube, LCSWa, Walter Meyer, MSb, J. Kathleen Tracy, PhDh

Departments of aPediatrics and bEpidemiology and Preventive Medicine, University of Maryland School of Medicine, Baltimore, Maryland

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

ABSTRACT

OBJECTIVES. The goals were to estimate the prevalence of parental depressive symptoms among parents at a pediatric primary care clinic and to evaluate the stability, sensitivity, specificity, and positive and negative predictive values of a very brief screen for parental depression.

METHODS. A total of 216 mothers (because 96% of caregivers were mothers, we use this term) bringing in children <6 years of age for child health supervision completed a parent screening questionnaire in a primary care clinic. The parent screening questionnaire, a brief screen for psychosocial problems developed for the study, includes 2 questions on depressive symptoms. Mothers then completed the computerized study protocol within 2 months. This included the parent screening questionnaire as well as the Beck Depression Inventory II. Different combinations of the depression questions were evaluated against Beck Depression Inventory II clinical cutoff values.

RESULTS. Twelve percent of the mothers met the Beck Depression Inventory II clinical cutoff value for at least moderate depressive symptoms. There was moderate stability of the screening questions. When a positive response to either or both of the 2 questions was considered, the sensitivity was 74%, the specificity was 80%, the positive predictive value was 36%, and the negative predictive value was 95%.

CONCLUSIONS. Maternal depressive symptoms are prevalent. A very brief screen can identify reasonably those who could benefit from additional evaluation and possible treatment. This should benefit mothers, families, and children.
Depression in Adults is a highly prevalent problem. Point prevalence rates of depression or depressive symptoms in women of parenting age range from 12% to 48%. As many as 23% of mothers bringing their children to a pediatric primary care clinic screened positively for depression.

The harmful impact of maternal depression on children’s health, development, and behavior has been amply demonstrated. Higher rates of low birth weight, poor growth, behavior and sleep problems, somatic complaints, learning difficulties, noninflicted injuries, and affective illness have been found for children of depressed mothers, compared with children whose mothers were not depressed. Field described several physiologic findings for infants of depressed mothers, including elevated norepinephrine and cortisol levels, lower vagal tone, and neurologic delays. McLennan and Kotelchuck found that mothers who reported high levels of depressive symptoms were less likely to implement preventive practices, such as using car seats and covering electrical outlets. Maternal depressive symptoms have also been associated with increased risk of infant hospitalization, use of corporal punishment, and lower likelihood of having a smoke alarm and using the supine sleep position. Minkovitz et al found that maternal depressive symptoms were associated with increased use of acute care, including emergency department visits, and decreased receipt of preventive care, including immunizations. Maternal depression has also been linked to diminished parenting expectations, confidence, and skills, as well as to child maltreatment.

Treatment of depression can be quite effective. For example, pharmacotherapy and psychotherapy are recommended for moderate to severe depression. Some studies have found medications alone to help patients with major depressive disorder significantly. Cognitive, behavioral, and interpersonal approaches have been found to be as effective as medication in treating mild to moderate depression, and psychotherapy alone is recommended. Recently, treatment of depressed mothers was shown to benefit their children. Another study found that mothers treated successfully for depression reported fewer behavior problems in their children.

Despite the high prevalence of depression and the effectiveness of treatment, few mothers are treated for depression, partly because of their lack of contact with health or mental health care services. Frequent visits with a pediatrician offer an opportunity to address this barrier. Olson et al reported that 40% of women who screened positive for depression in pediatric clinics accepted referrals for additional intervention.

On the basis of the aforementioned evidence, pediatricians have been encouraged to include screening for depression in primary care settings. In addition, mothers seem interested in pediatricians doing so. Nevertheless, pediatricians have been found to identify depression rarely. Few have implemented a systematic approach to screening for maternal depression.

The US Preventive Services Task Force has recommended routine screening of all adults for depression, with the use of 2 questions pertaining to mood and anhedonia (ie, loss of pleasure). Wholey et al found that 2 questions performed as well as longer screening measures, such as the Beck Depression Inventory (BDI) and the Center for Epidemiological Studies-Depression scale. In addition, both interview and paper-and-pencil approaches have been validated, although Olson et al found a much higher rate of positive results with paper screening (22.9% vs 7.6%).

Despite considerable evidence supporting screening of mothers for depression in pediatric primary care, we are aware of only a single study that examined the validity of brief screening tools in such settings. This study compared a 3-item version with the 8-item Rand measure from which it was derived. The objective of the present study was to test the stability, sensitivity, specificity, and positive and negative predictive values of a 2-item screen for parental depression in a pediatric primary care clinic. This was performed as part of a larger project that trained residents to address prevalent psychosocial problems, implemented a parent screening questionnaire (PSQ) for the targeted problems in their continuity clinics, and included a social worker in efforts to meet identified needs.

METHODS
Participants

The current study was part of a research project with a quasi-experimental design that was based in a university-affiliated, pediatric, resident, continuity clinic. After stratification for medicine/pediatrics residents, 2 clinic days were chosen randomly for the intervention, with routine care provided on the other clinic days. The intervention involved asking parents to complete a PSQ for identification of possible parental risk factors for child maltreatment. Pediatric residents assigned to intervention clinic days were trained to assess briefly, to provide initial management for, and consider referral of parents who answered yes to any of the screening questions. Parents bringing their children (<6 years of age) for a regular checkup visit on intervention clinic days were eligible to participate in the component of the research that involved validation of the PSQ. Parents needed to speak English and to agree to participate in the study. Exclusionary criteria included parents who did not speak English, parents with another child in the study, and families whose child receiving the checkup examination was in foster care.

Five hundred seven parents were approached to participate; some left the clinic before a research assistant could speak with them. Of those approached, 382 (75%)
agreed to participate, and 308 (81% of those recruited) completed the initial study protocol (Fig 1). The remaining 74 of those recruited did not keep the scheduled appointments, and a few actively withdrew. Of the 308, 92 did not complete the PSQ in the clinic within a 2-month period or had incomplete data. There were thus 92 who needed to be excluded, leaving 216 for the analyses.

We compared the 216 subjects with complete data with the 92 subjects who completed the PSQ either outside the 2-month window or not at all. There were no differences in terms of the children’s age, race, and gender. There were also no differences regarding the parents’ age, gender, educational status, and marital status, although there were more employed mothers in the group with complete data, compared with those missing the PSQ (36% vs 24%; P = .05). The groups had similar average numbers of adults and children in the home, and most were receiving Medical Assistance. The 2 groups did not differ on the PSQ questions (on the computerized protocol) pertaining to depression. We also compared the PSQ responses of the 216 subjects with those of all 548 parents who completed the PSQ but were not in the final sample (764 – 216 parents). Those in the study sample were more likely to answer yes to one of the depression screens (27% vs 20%; P = .04).

The present article is based on the sample of 216 caregivers in the intervention group who completed the PSQ in the primary care clinic and a computerized version in the study protocol within 2 months (Table 1). Most of the caregivers were single mothers who were unemployed. They averaged 25 years of age and had varying levels of education. The children averaged 11.8 months of age. Approximately one half were male, and most were black. Families had an average of 2.3 children and 2.2 adults in the home, and most were receiving Medical Assistance.

**Procedure**

Parents bringing their <6-year-old children to the pediatric primary care clinic to see a resident for child health supervision were given the PSQ to complete, voluntarily, while waiting for their appointment. Only parents assigned to intervention clinic days received the PSQ. Parents with >1 child to be seen were asked to focus on the youngest one. They were informed that these questionnaires would be used for research purposes, and they were asked to sign their consent on the back of the form. When completed, the questionnaires were given to the continuity clinic resident who had been trained to address the targeted problems.

In addition, parents were approached and asked to participate in research to validate the PSQ. If they were interested, then informed consent procedures were implemented, as approved by our university’s human subjects review committee. Parents were asked to return within 2 weeks to complete an audio-computerized self-interview of the study protocol that included the PSQ. The median time before this was completed was 8 days. The computerized interview also had more comprehensive, standardized measures of the targeted problems, such as the BDI II, to examine the validity of the screen. Subjects were paid $60 for completing the study protocol.

**Measures**

**PSQ**

The PSQ was developed as a tool for pediatric practices to screen parents briefly for major psychosocial problems and risk factors for child maltreatment (eg, maternal depression, substance abuse in the family, or intimate partner violence). We reviewed the literature to prioritize risk factors that were reasonably amenable to intervention. Next, validated screening measures for several risk factors were reviewed and incorporated. The 2-item depression screen developed by Whooley et al was found to have very good psychometric properties. However, this measure required minor modifications in the wording to be suitable for a low-income, less-educated population. Consequently, the words “bothered by” were removed from the questions, “During the past month, have you often been bothered by feeling down, depressed, or hopeless?” and “During the past month, have you often been bothered by little interest or pleasure in doing things?” An advisory committee of community pediatricians guided the development of the PSQ and recommended keeping the PSQ brief, with a yes/no response set. The PSQ was designed to be easy for pediatricians to assess by readily noting the yes responses. The PSQ was then pilot tested with ~40 parents in neighboring pediatric clinics. Parents’ input was used to reduce the number of items and to clarify those that were unclear. Because the PSQ solicits sensitive information, it was important to
frame this in a supportive context. Therefore, the screen started with an introduction that conveyed an empathetic tone, concern about safety, and a willingness to help (Appendix).

The 2 screening questions were analyzed both separately and in combination. There were 2 combinations. If parents responded yes to either question, then the screen was considered to be positive (ie, the problem may be present). A screen with responses of no to both questions was coded as a negative screen. In the alternative approach, a positive screen was considered only when there were yes responses to both questions; otherwise, the screen was considered to be negative.

**BDI II**

The BDI II is a widely used measure for screening for depression in adults. It has high internal consistency (ranging from 0.86 to 0.93) and high test-retest reliability. It has excellent concurrent validity with other self-report measures and clinical ratings of depression. There is also substantial evidence of its construct validity. Its reliability, validity, and factor structure are similar across diverse populations. In addition, it has been found to have adequate positive and negative predictive values. Although the BDI II does not provide a definitive diagnosis of depression, it does identify the presence and severity of symptoms consistent with DSM-IV criteria. Ideally we would have conducted clinical interviews to establish a diagnosis of depression, but resources for the study precluded this option. The BDI II thus served as a reasonable compromise to validate a very brief, 2-item screen.

The BDI II is composed of 21 questions, each with scores ranging from 0 to 3, reflecting the frequency with which the problem occurs. The total BDI II score (range: 0–63) is the sum of the scores for the 21 questions. A total BDI II score of >19 has been found to correspond to moderate or severe clinical depression. Missing responses for any of the questions make it difficult to ascertain an accurate total BDI II score. To address this problem, any parent with a total BDI II score of >19, regardless of missing responses, was considered at risk for moderate depression. In addition, if the total BDI II score and number of missing responses were in the following combinations, then parents were categorized as not clinically depressed: score of ≤16 and 1 missing item (a score of 3 for the missing item could yield a total score of 19, just short of the clinical cutoff point), score of ≤13 and 2 missing items, score of ≤10 and 3 missing items, score of ≤7 and 4 missing items, score of ≤4 and 5 missing items, or score of ≤1 and 6 missing items. Twenty participants were excluded because missing items could have placed them over the clinical cutoff point; therefore, the final sample size was reduced to 196.

### Data Analysis

**Stability**

To determine the stability of the depression screen, we compared the same 2 depression questions on the clinic PSQ and the computerized PSQ. Because a caregiver could have completed >1 clinic PSQ, the last one within

| TABLE 1 Demographic Characteristics of Study Participants (N = 216) |
|------------------|------------------|------------------|
| **Proportion, %** | **Mean ± SD**    | **Median**       |
| Caregiver        |                  |                  |
| Relationship to child | 94               |                  |
| Mother           | 94               |                  |
| Father           | 4                |                  |
| Marital status   |                  |                  |
| Married          | 8                |                  |
| Single           | 87               |                  |
| Age, y           | 25.2 ± 6.8       | 24.0             |
| Education        |                  |                  |
| Less than high school | 35             |                  |
| High school or GED | 39              |                  |
| At least some college | 26           |                  |
| Employed, part-time or full-time | 36         |                  |
| Child            |                  |                  |
| Age, mo          | 11.8 ± 14.1      | 5.0              |
| Male gender      | 56               |                  |
| Race             |                  |                  |
| Black            | 92               |                  |
| White            | 3                |                  |
| Mixed            | 5                |                  |
| Family           |                  |                  |
| No. of children in home | 2.3   | 2.0              |
| No. of adults in home | 2.2    | 2.0              |
| Receiving Medical Assistance | 93 |                  |

GED indicates general equivalency diploma.
When the clinic PSQ depression screen was compared with the computerized version, \( \kappa \) values of \( 0.62, 0.45, 0.57, \) and \( 0.50 \) were observed for questions A, B, A or B (either one yes), and A and B (both yes), respectively.

### Validity

To determine the validity of the clinic depression screen, we compared it against the BDI II by examining the sensitivity, specificity, positive predictive values, and negative predictive values. We compared each screening item, and the 2 items in different combinations, with the BDI II assessment (depressed/not depressed) by using the cutoff point of \( >19 \), representing moderate to severe depression.

### RESULTS

#### Prevalence Rates

The prevalence rates for depression based on the clinic PSQ and the BDI II are presented in Table 2. The 2 screening questions individually yielded similar rates (17%–19%), close to that found with the BDI II. However, the rate was considerably higher (27%) when a yes to either or both screening questions (A or B) was considered a positive result and considerably lower (9%) when only a yes to both screening questions (A and B) was considered a positive result.

#### Stability

When the clinic PSQ depression screen was compared with the computerized version, \( \kappa \) values of \( 0.62, 0.45, 0.57, \) and \( 0.50 \) were observed for questions A, B, A or B (either one yes), and A and B (both yes), respectively.

#### Validity

Table 3 presents results pertaining to the validity of the depression screen items, compared with the BDI II. The sensitivity of the screen was maximized when a positive screen was defined as a yes response to either or both of the depression questions (A or B). Use of this definition also maximized the negative predictive value; a respondent with a negative screen had a 95% likelihood of not being depressed.

### DISCUSSION

This study confirms the high prevalence of depression or depressive symptoms among parents of young children reported by others, although the rate found in this study was on the low side (22%, including mild depressive symptoms). The study also demonstrates that a 2-item screen has adequate properties (stability and validity) for accurate identification of parents who may be moderately depressed and need at least additional evaluation. These findings support the work of others who found brief screens for depression to be as effective as longer screens in adult clinics. Kemper and Babonis reported a sensitivity of 100%, a specificity of 88%, and a positive predictive value of 66%. It should be noted, however, that they compared a 3-item version with the 8-item Rand measure from which it was derived.

High sensitivity is one goal of a screening questionnaire, to ensure that few persons with the condition are missed. The 74% sensitivity obtained in this study is satisfactory, although not as high as we hoped. Nevertheless, screening for depression helps identify many parents in need of help, albeit not all, and seems preferable to not screening at all. It is noteworthy that depression is masked frequently and may not be apparent in brief encounters. Another consideration is that those who do not disclose a problem may not be at a stage where they are ready to address it. Therefore, although high sensitivity is the goal, it may be that those who are missed are least likely to engage in treatment. It is possible that a seed may be sown simply through questioning and parents may appreciate pediatricians’ interest in the parents’ own health. Over time, parents may learn to recognize pediatricians as a resource and may confide more in them. It seems that pediatricians generally have good rapport with parents, and many parents are comfortable discussing their problems with pediatricians. This should be helpful in facilitating the receipt of services parents may need. Indeed, parents may be grateful for the attention shown to them and may be more satisfied with their children’s pediatric care.

There is also a concern regarding false-positive results, unduly “labeling” a parent as depressed. Indeed, the 36% positive predictive value with use of a positive response to either question indicates that only one third of those who screened positive had moderate or worse depressive symptoms. It is critical to distinguish between a screen and a diagnosis. Pediatricians and parents need to appreciate the difference; the screen only identifies a possible problem. Additional evaluation is clearly warranted. Pediatricians can be taught to assess parental depression briefly and to clarify whether a mental health evaluation seems indicated. Those working closely with
TABLE 3  Sensitivity, Specificity, and Predictive Values for the PSQ Depression Items Compared With Total BDI II Scores of $>$ 19

<table>
<thead>
<tr>
<th>Depression Screen Item</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Positive Predictive Value, %</th>
<th>Negative Predictive Value, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. “Lately, do you feel down, depressed, or hopeless?” (yes)</td>
<td>63</td>
<td>87</td>
<td>43</td>
<td>94</td>
</tr>
<tr>
<td>B. “During the past month, have you felt very little interest or pleasure in the things you used to enjoy?” (yes)</td>
<td>44</td>
<td>87</td>
<td>34</td>
<td>91</td>
</tr>
<tr>
<td>A or B (yes to either question)</td>
<td>74</td>
<td>80</td>
<td>36</td>
<td>95</td>
</tr>
<tr>
<td>A and B (yes to both questions)</td>
<td>33</td>
<td>95</td>
<td>47</td>
<td>91</td>
</tr>
</tbody>
</table>

BDI II scores of $>$ 19 suggest moderate or severe depression.

social workers or other mental health professionals may be able to assess the situation readily in the practice setting. Alternatively, a referral to a community mental health resource may be needed. In these ways, concerns regarding possible erroneous labeling can be minimized.

In order to comprehensively assess a screening test’s performance, one needs to examine its sensitivity, specificity, and positive and negative predictive values. There are inevitable compromises depending where one sets the threshold for “positive.” When positive was defined as a yes to either screening question, the rather high specificity indicates many true-negatives and few false-positives. The 74% sensitivity indicates that three quarters of those with depressive symptoms were detected (true-positives), but one quarter were missed (false-negatives). The 95% negative predictive value adds that most of those who screened negative were identified correctly as not being depressed.

It is likely that, for some parents, inquiries into their own health in a pediatric clinic may be surprising and may seem intrusive. In the course of this study, however, we were not aware of any objections to the screening questions. Responses were voluntary and parents could choose, for a variety of reasons, not to disclose their problems. A study found that mothers were more likely to discuss depressive symptoms with family members and friends than with a pediatrician, because of mistrust and fear of being judged, although they were receptive to doing so if they thought that the pediatrician “knew them well.”

In addition to training and commitment, pediatricians need valid practical tools. The depression screen in this study was included in a brief PSQ aimed at identifying several psychosocial problems. The use of such questionnaires has been found to be useful and efficient in pediatric practice. The questionnaires can be filled out by parents while they are waiting for their child to be seen. The pediatrician can peruse the questionnaire quickly and focus on problem items.

Several questions remain for future research. What are the best methods for screening? One study found that more mothers reported depressive symptoms on a paper-based screen, compared with the same questions in a scripted interview (22.9% vs 5.7%). How can sensitive new questions be introduced into the well-child visit? We began the PSQ by explaining that we were addressing problems such as parental depression requires some time. Pediatricians can be trained to do this efficiently, with a few key questions. With practice, pediatricians likely can become competent and comfortable addressing such problems. Our training included brief specific information to access community resources, as well as parent handouts. A larger issue is raised, however. If the time available for these visits remains so limited, then there is a need to reconsider how best to take advantage of the opportunity and to prioritize areas in which pediatricians can be most helpful and can make an important difference in children’s health and well-being. This requires rethinking some aspects of pediatric primary care that have become ritualized but may have little value.

In addition to training and commitment, pediatricians need valid practical tools. The depression screen in this study was included in a brief PSQ aimed at identifying several psychosocial problems. The use of such questionnaires has been found to be useful and efficient in pediatric practice. The questionnaires can be filled out by parents while they are waiting for their child to be seen. The pediatrician can peruse the questionnaire quickly and focus on problem items.

Several questions remain for future research. What are the best methods for screening? One study found that more mothers reported depressive symptoms on a paper-based screen, compared with the same questions in a scripted interview (22.9% vs 5.7%). How can sensitive new questions be introduced into the well-child visit? We began the PSQ by explaining that we were addressing problems such as parental depression requires some time. Pediatricians can be trained to do this efficiently, with a few key questions. With practice, pediatricians likely can become competent and comfortable addressing such problems. Our training included brief specific information to access community resources, as well as parent handouts. A larger issue is raised, however. If the time available for these visits remains so limited, then there is a need to reconsider how best to take advantage of the opportunity and to prioritize areas in which pediatricians can be most helpful and can make an important difference in children’s health and well-being. This requires rethinking some aspects of pediatric primary care that have become ritualized but may have little value.

In addition to training and commitment, pediatricians need valid practical tools. The depression screen in this study was included in a brief PSQ aimed at identifying several psychosocial problems. The use of such questionnaires has been found to be useful and efficient in pediatric practice. The questionnaires can be filled out by parents while they are waiting for their child to be seen. The pediatrician can peruse the questionnaire quickly and focus on problem items.

Several questions remain for future research. What are the best methods for screening? One study found that more mothers reported depressive symptoms on a paper-based screen, compared with the same questions in a scripted interview (22.9% vs 5.7%). How can sensitive new questions be introduced into the well-child visit? We began the PSQ by explaining that we were addressing problems such as parental depression requires some time. Pediatricians can be trained to do this efficiently, with a few key questions. With practice, pediatricians likely can become competent and comfortable addressing such problems. Our training included brief specific information to access community resources, as well as parent handouts. A larger issue is raised, however. If the time available for these visits remains so limited, then there is a need to reconsider how best to take advantage of the opportunity and to prioritize areas in which pediatricians can be most helpful and can make an important difference in children’s health and well-being. This requires rethinking some aspects of pediatric primary care that have become ritualized but may have little value.
seems reasonable that these likely would be good candidates to identify for additional evaluation and possible treatment. When we examined the sensitivity by using a lower threshold that included mild depression (score of >13), the sensitivity was only 62%.

What are the best questions to ask? The questions in the present study worked well and are similar to those recommended by the US Preventive Services Task Force, but additional research is needed, and it is likely that wording may vary slightly to be culturally sensitive for different populations. What are the best times to ask these questions? We planned to do so only periodically, but for logistic reasons it was easiest to have parents of preschool-aged children complete the PSQ at each well-child visit. There were no complaints, but it may be preferable to ask these questions less frequently or to intersperse these questions with others addressing different issues. Most importantly, there is a need to study whether screening leads to parents in need receiving appropriate services and experiencing improvement in their depression. We are examining this issue currently.

There are signs that pediatricians may be ready to embrace some changes to practice. Olson et al found that 4 community pediatric practices were able to incorporate screening for depression in well-child visits. In a survey of 508 community pediatricians, 57% felt responsible for recognizing maternal depression and 21% were willing to modify their approach to identification. Still, the survey also found that inadequate training (65%) and time to assess the problem (70%) were major barriers. Training is clearly needed to help pediatricians address parental depression. Still, Olson et al found that 43% of pediatricians did not feel responsible for identifying maternal depression; more needs to be known about their resistance, and education is needed to demonstrate the importance of this problem.

Our study has several limitations. It was conducted in an urban, low-income, mostly black population, necessitating caution about applicability to other groups. Even within our clinic, we found that those who participated were more likely to report depressive symptoms than those not in the study. The sensitivity and specificity should be similar in other settings. The predictive values, however, may vary in other populations.

The PSQ was first administered in a primary care clinic and then repeated with the study protocol in a self-administered computerized interview. The variation in methods could contribute to differing responses and precludes strict examination of test-retest reliability. Despite this, the screening questions proved to be quite stable over time.

Ideally, the screen would have been validated against a standard diagnostic measure such as a comprehensive clinical interview. We lacked the resources to do so, relying instead on the BDI II, a well-standardized, lengthier screen for depression. It should be noted that this measure has performed well in comparison with clinical interviews. In addition, Field et al found that depressive symptoms, in contrast to a diagnosis of depression, were more frequently related to problematic mother-child interactions.

CONCLUSIONS

The high prevalence of parental depression is clear, and the harmful impact on children has been amply demonstrated. Identifying and addressing maternal depression, thereby promoting the health of mothers, should decrease significantly the substantial morbidity in children. This study provides a useful first step in showing how a very brief screen can identify accurately parents who may be depressed and in need of help.

ACKNOWLEDGMENT

This research was supported by grant 90-CA-1695 from the Office on Child Abuse and Neglect, Administration for Children, Youth, and Families.

REFERENCES


19. Weissman M. Recent non-medication trials of interpersonal psychotherapy for depression. *Int J Neuropsychopharmacol.* 2006;20:1–6


Parent Screening Questionnaire
A Safe Environment for Every Kid (SEEK)

Dear parent or caregiver: Being a parent is not easy. We want to help families have a safe environment for kids. We are asking everyone these questions. Please answer the questions about your child being seen today for a check-up. They are about issues that affect many families. If there’s a problem, we’ll try to help.

Today’s Date: ___/___/200__
Child’s Date of Birth: ___/___/____
Sex of Child: □ Male □ Female

PLEASE CHECK
□ Yes □ No Do you need the telephone number for Poison Control?
□ Yes □ No Do you need a smoke alarm for your home?
□ Yes □ No Does anyone smoke tobacco at home?
□ Yes □ No Is there a gun in your home?
□ Yes □ No In the last year, did you worry that your food would run out before you got money or food stamps to buy more?
□ Yes □ No Do you worry that your child may have been physically abused?
□ Yes □ No Do you worry that your child may have been sexually abused?
□ Yes □ No Lately, do you often feel down, depressed, or hopeless?
□ Yes □ No Do you often feel lonely?
□ Yes □ No During the past month, have you felt little interest or pleasure in the things you used to enjoy?
□ Yes □ No Do you often feel your child is difficult to take care of?
□ Yes □ No Do you wish you had more help with your child?
□ Yes □ No Do you feel so stressed you can’t take another day?
□ Yes □ No Do you sometimes find you need to hit/spank your child?
□ Yes □ No In the past year, have you or your partner had a problem with drugs or alcohol?
□ Yes □ No In the past year, have you or your partner felt the need to cut back on alcohol?
□ Yes □ No Have you ever been in a relationship in which you were physically hurt or threatened by a partner?
□ Yes □ No In the past year, have you been afraid of a partner?
□ Yes □ No In the past year have you thought of getting a court order for protection?
□ Yes □ No Are there any problems you’d like help with today?

Please give this form to the doctor or nurse you’re seeing today. Thank you.
Factors Associated With Identification and Management of Maternal Depression by Pediatricians

Amy M. Heneghan, MD, Linda H. Chaudron, MD, MS, Amy Storfer-Isser, MD, Elyse R. Park, PhD, Kelly J. Kelleher, MD, MPH, Ruth E. K. Stein, MD, Kimberly Eaton Hoagwood, PhD, Karen G. O’Connor, BS, Sarah McCue Horwitz, PhD

Departments of Pediatrics and Epidemiology and Biostatistics, School of Medicine, Case Western Reserve University, Cleveland, Ohio; Departments of Psychiatry, and Obstetrics and Gynecology, University of Rochester, Rochester, New York; Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts; Department of Pediatrics, College of Medicine and Public Health, Ohio State University, Columbus, Ohio; Department of Pediatrics, Albert Einstein College of Medicine, Children’s Hospital at Montefiore, Bronx, New York; Department of Psychiatry, Columbia University, New York, New York; Department of Research, American Academy of Pediatrics, Elk Grove Village, Illinois

Financial Disclosure: Dr Chaudron discloses grant funding from Forest Laboratories and the Wyeth Pharmaceuticals speaker’s bureau.

ABSTRACT

OBJECTIVE. We sought to identify characteristics of pediatricians that were associated with identification or management (referral and/or treatment) of mothers with depression.

METHODS. A cross-sectional survey was mailed to a random sample of 1600 of the 50 818 US nonretired members of the American Academy of Pediatrics. Overall, 832 responded, with 745 responses from nontrainee members. The 662 fellow nontrainee members who engaged in direct patient care and completed information on identifying, referring, and treating maternal depression were included in the analyses.

RESULTS. A total of 511 of 662 respondents reported identifying maternal depression; of those who reported identifying maternal depression, 421 indicated they referred and 29 that they treated maternal depression in their practices. Pediatricians who are older, work in practices that provide child mental health services, see primarily white patients, use 1 method to address maternal depression, agree that pediatricians should be responsible for identifying maternal depression, think that maternal depression has an extreme effect on children’s mental health, and are attitudinally more inclined to identify or manage maternal depression had significantly higher odds of reporting identification of maternal depression. Positive correlates of identification and management of maternal depression included practicing in the Midwest, using ≥1 method to address maternal depression, working in a practice that provides child mental health services, thinking that caregiving problems attributable to maternal health have an extreme effect on children’s physical health, having attitudes that are more inclined to identify and to manage maternal depression, and usually inquiring about symptoms routinely to identify maternal depression.

CONCLUSIONS. Pediatricians’ practice characteristics and attitudes are associated with their identification and management of mothers with depression.
M A T E R N A L  D E P R E S S I O N. A global term encompassing a range of depressive symptoms and syndromes, may affect women immediately after childbirth or later in the child-rearing years.1,2 For women 15 to 44 years of age, depression is the leading cause of worldwide disease burden.3 Eight to 12% of women experience postpartum depression,4,5 although twice that number of women may experience elevated depressive symptoms.6,7 The risk of depression and depressive symptoms is increased in women with young children, women with >1 child, and women of low socioeconomic status.8–14

Negative consequences for women with depression include increased risk of future depression, impaired mother-child bonding, and even overt thoughts of harming their infant.15–22 In addition, extensive research has shown that maternal depression experienced during the postpartum period and beyond can have a negative impact on children’s social, cognitive, and behavioral development.1,23–30

Depressed women are likely to be overlooked because of their pattern of medical care use.31–33 Although typically pediatricians do not treat the mothers of their young patients, they are the health care professionals with the most frequent contact with depressed women of childbearing age.33 The American Academy of Pediatrics (AAP) recommends a minimum of 7 well-child visits for children during the first 1 year of life.34 Therefore, pediatricians are in an excellent position to discuss maternal depression during well-child visits. Bright Futures: Guidelines for Health Supervision encourages pediatricians to support families as part of providing primary health care for children.35 In particular, Bright Futures in Practice: Mental Health recommends that, during health supervision visits, pediatricians “ask parents about any new experiences or stresses in their own lives, and about feelings of sadness, sleep problems, loss of interest in activities they used to enjoy, and other specific symptoms of depression.”36 The AAP Task Force on the Family has advocated for “family-oriented” pediatric care, to improve family outcomes.37 The task force has stated specifically that the “health and well-being of children is inextricably linked to their parents’ physical, emotional, and social health.”38 In addition, the Future of Pediatrics Education II residency education guidelines have called specifically for educating residents about family health and maternal depression.39

Despite a growing orientation toward addressing family and maternal issues in pediatric practice, pediatricians do not address these issues readily. Heneghan et al40 found that pediatricians recognized only one fourth of mothers with depressive symptoms. Chaudron et al41 showed that, even when well-structured screening programs were incorporated into pediatric primary care practice, only one half of mothers were screened for depression. In a national sample of practicing pediatricians, Olson et al42 described pediatricians’ perceived roles, management practices, and barriers encountered in addressing maternal depression. Fifty-seven percent thought that it was their responsibility to recognize maternal depression, but only 7% felt responsible for providing treatment. Time constraints and incomplete training were barriers most often noted by the pediatricians surveyed. More-recent work by Horwitz et al43 demonstrated that pediatricians endorsed similar barriers to identification and management of maternal depression.

Physician and practice characteristics that are associated with increased identification of depressive symptoms have been examined in adult settings,44–51 but no study has examined the factors that influence pediatricians’ identification and management of maternal depression. Because pediatricians are uniquely poised to assist mothers with depression, we sought to identify characteristics of pediatricians that are associated with identification and management, in the form of referral or treatment, of mothers with depression. In particular, we examined pediatricians’ sociodemographic and practice characteristics, attitudes, and practices associated with 2 outcomes, namely, (1) pediatricians who have identified maternal depression, compared with pediatricians who have not, and (2) pediatricians who have identified and managed maternal depression, compared with pediatricians who have only identified maternal depression.

METHODS

Survey Administration

The study population for this research consisted of the 50,818 US nonretired members of the AAP. The AAP has conducted a periodic survey 3 or 4 times each year since 1987, to provide information on current pediatric topics to inform policy, to develop new initiatives, or to evaluate current projects.52 Using the AAP periodic survey initiative, an AAP researcher (Ms O’Connor) selected and mailed the survey to 1600 members beginning in March 2004, with a sixth and final mailing in August 2004. The questionnaire was 8 pages in length, contained largely closed-ended questions, had been pretested, and was approved by the AAP institutional review board before the initial mailing. Overall, 832 AAP members (52%) responded. Fifty-seven percent of nontrainee members (ie, not residents or fellows) responded. Of those, 662 nontrainee members who engaged in direct patient care and completed the information on identifying, referring, and treating maternal depression were included in the analyses.

Survey Questionnaire

Respondents were asked a broad range of sociodemographic, practice characteristic, and patient characteristic questions that were used in previous periodic surveys. Pediatricians’ impressions about the prevalence of maternal depression and caregiving problems attrib-
utable to a mother’s physical or mental health issues were ascertained. The impact of maternal psychosocial problems on children’s mental and physical health was rated by pediatricians as little/none, moderate, or extreme/great. Pediatricians were asked, “Do you or others working in your practice (eg, other physicians, nurse practitioner, social worker, psychologist, counselor/therapist, etc) provide medication, counseling, or psychotherapy for mental health problems for children or adolescents?” Pediatricians were asked, “During typical pediatric visits have you, yourself, ever identified depression in a mother of a patient?” Those who had identified mothers with depression were asked whether they had ever referred a mother for diagnosis or treatment or treated a mother (ie, provided in-depth counseling and/or medications). The term “manage,” as used throughout the article, was constructed to include pediatricians who have ever referred and/or treated mothers for depression. In addition, the frequency of use of specific passive and active methods to identify maternal depression was assessed as a 3-level response (usually, sometimes, or never). These methods ranged from observation to use of a screening tool. Pediatricians were asked whether they used currently, or would like to use, any of 12 different methods to address maternal depression. Examples included a team approach with gynecologists, mental health professionals, and pediatricians; partnerships with schools and early childhood providers; bulletin boards with information and resources in the waiting room or examination rooms; and use of Bright Futures in Practice: Mental Health. To assess attitudes about identifying, referring, and treating maternal depression, pediatricians were asked whether they agreed, were neutral, or disagreed that pediatricians should be responsible for each of these 3 actions. Additional assessment of pediatricians’ attitudes toward maternal depression was performed by using 3 scales developed by Park et al (E.R.P., A.S.-I., K.J.K., et al, unpublished data, 2006). These scales are based on previous work by Williams et al, Olson et al, Park et al,55 and McLennan et al.56 They measure the following attitudinal attributes of pediatricians: acknowledgment of maternal depression (3 items), perceptions of mothers’ beliefs (3 items), and treatment of maternal depression (4 items). Reliability ratings (Cronbach’s α) are .59, .72, and .79, respectively. Higher scores correspond to more-favorable attitudes about managing maternal depression.

Sample Weights

Characteristics from the AAP member list for responders and nonresponders for the overall sample and the nontrainee members were reported elsewhere.43 Bivariate comparisons of nontrainee responders and nonresponders surveyed showed that women, pediatricians <40 years of age, and fellows and candidate fellows were significantly more likely to respond. To avoid potential bias created by different survey nonresponse rates, poststratification sample weights were created and weighted analyses were performed. The sample weights were created with a saturated logistic regression model with age group (≥40 vs <40 years), gender, and the 2-way interaction between age and gender as predictors of response and the weights as the inverse of the probability of response. The weights were rescaled such that the mean was unity and the sum of the weights was equal to the sample size. These procedures ensured that the sample was representative of the gender and age distribution of AAP membership.

Statistical Analyses

Weighted means and SDs and counts with weighted percentages were used to summarize sample characteristics for continuous and categorical data, respectively. Bivariate and regression analyses were performed for each of the 2 outcomes (has identified versus has not identified and has identified and managed versus has identified only) by using the same analytic approach. The Rao-Scott χ² test, weighted logistic regression, and weighted linear regression were used to assess bivariate associations between each covariate and each outcome. Covariates with a statistically significant (P < .05) association with the outcome were further assessed in multivariate, weighted, logistic regression models. Main-effects models were fitted, and statistically significant predictors were retained. On the basis of the results of the main-effects models, 2-way interactions of clinical relevance were evaluated. The results are summarized with adjusted odds ratios (ORs) and 95% confidence intervals (CIs). Analyses were performed by using procedures for survey data in SAS 9.1 (SAS Institute, Cary, NC).

RESULTS

Figure 1 depicts the rates of identification, referral, and treatment of maternal depression among pediatric providers. Of the 745 nontrainee respondents, 662 engaged in direct patient care and completed questions regarding identification and management of maternal depression. Of those 662 respondents, 511 (77%) reported ever identifying maternal depression; of those who reported identifying, 421 (82%) also indicated that they referred maternal depression in their practices. Of those, 363 (86%) reported the number of mothers referred for maternal depression in the past year (median: 4 mothers; range: 0–50 mothers). Only 29 pediatricians (7% of those who identify mothers with depression, or 4% of the entire sample) reported ever treating a mother for maternal depression.

Table 1 displays the weighted physician and practice characteristics for the 662 nontrainee respondents, stratified according to each outcome. Compared with pedia-
FIGURE 1
Respondents who have ever identified, referred, or treated maternal depression. DPC indicates direct patient care.

TABLE 1 Pediatrician and Practice Characteristics

<table>
<thead>
<tr>
<th>Does Not Identify</th>
<th>Identifies</th>
<th>P</th>
<th>Identifies Only</th>
<th>Identifies and Manages*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N = 151)</td>
<td>(N = 511)</td>
<td></td>
<td>(N = 90)</td>
<td>(N = 421)</td>
</tr>
<tr>
<td><strong>Pediatrician’s gender, (n) (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>66 (49.9)</td>
<td>209 (46.8)</td>
<td>43 (53.8)</td>
<td>166 (45.2)</td>
</tr>
<tr>
<td>Female</td>
<td>85 (50.1)</td>
<td>302 (53.2)</td>
<td>5077</td>
<td>255 (54.8)</td>
</tr>
<tr>
<td><strong>Pediatrician’s age, mean ± SD, y</strong></td>
<td>42.3 ± 10.6</td>
<td>46.2 ± 10.3 &lt;0.0001</td>
<td>45.4 ± 10.5</td>
<td>46.4 ± 10.2 &lt;0.0001</td>
</tr>
<tr>
<td><strong>Time in practice, mean ± SD, y</strong></td>
<td>10.4 ± 9.8</td>
<td>14.3 ± 10.5 &lt;0.0001</td>
<td>13.3 ± 11.4</td>
<td>14.5 ± 10.3 &lt;0.0001</td>
</tr>
<tr>
<td><strong>Pediatrician’s race, (n) (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>91 (60.6)</td>
<td>382 (75.1)</td>
<td>57 (64.6)</td>
<td>325 (77.4)</td>
</tr>
<tr>
<td>Asian</td>
<td>37 (24.5)</td>
<td>60 (11.6)</td>
<td>15 (15.9)</td>
<td>45 (10.6)</td>
</tr>
<tr>
<td>All others</td>
<td>23 (14.9)</td>
<td>69 (13.3)</td>
<td>0.002</td>
<td>51 (12.0)</td>
</tr>
<tr>
<td><strong>Location of practice, (n) (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>65 (44.6)</td>
<td>210 (42.9)</td>
<td>32 (37.0)</td>
<td>178 (44.1)</td>
</tr>
<tr>
<td>Suburban</td>
<td>65 (45.1)</td>
<td>208 (42.5)</td>
<td>39 (43.4)</td>
<td>169 (42.4)</td>
</tr>
<tr>
<td>Rural</td>
<td>15 (10.3)</td>
<td>70 (14.6)</td>
<td>4094</td>
<td>54 (13.5)</td>
</tr>
<tr>
<td><strong>Region of practice, (n) (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>36 (24.5)</td>
<td>145 (28.4)</td>
<td>29 (32.2)</td>
<td>116 (27.6)</td>
</tr>
<tr>
<td>Midwest</td>
<td>28 (18.6)</td>
<td>109 (21.2)</td>
<td>7 (7.5)</td>
<td>102 (24.2)</td>
</tr>
<tr>
<td>South</td>
<td>60 (38.7)</td>
<td>157 (30.9)</td>
<td>31 (35.0)</td>
<td>126 (30.0)</td>
</tr>
<tr>
<td>West</td>
<td>27 (18.2)</td>
<td>100 (19.5)</td>
<td>3563</td>
<td>77 (18.2)</td>
</tr>
<tr>
<td><strong>Patients have specific pediatrician, (n) (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>100 (65.6)</td>
<td>382 (74.8)</td>
<td>68 (75.6)</td>
<td>314 (74.7)</td>
</tr>
<tr>
<td>No</td>
<td>40 (27.1)</td>
<td>76 (14.6)</td>
<td>14 (15.4)</td>
<td>62 (14.5)</td>
</tr>
<tr>
<td>Do not know</td>
<td>11 (7.3)</td>
<td>53 (10.5)</td>
<td>0.0018</td>
<td>8 (9.0)</td>
</tr>
<tr>
<td>≥80% of patients have private health insurance, (n) (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>42 (27.3)</td>
<td>154 (30.4)</td>
<td>28 (30.4)</td>
<td>126 (30.4)</td>
</tr>
<tr>
<td>No</td>
<td>78 (51.5)</td>
<td>272 (53.3)</td>
<td>45 (51.5)</td>
<td>227 (53.7)</td>
</tr>
<tr>
<td>Do not know</td>
<td>31 (21.2)</td>
<td>85 (16.3)</td>
<td>3769</td>
<td>68 (15.9)</td>
</tr>
<tr>
<td><strong>Patients’ race, (n) (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;75% white</td>
<td>121 (80.5)</td>
<td>328 (65.8)</td>
<td>51 (59.0)</td>
<td>277 (67.2)</td>
</tr>
<tr>
<td>≥75% white</td>
<td>28 (19.5)</td>
<td>169 (34.2)</td>
<td>0.008</td>
<td>134 (32.8)</td>
</tr>
<tr>
<td><strong>No. of ambulatory visits per wk, mean ± SD</strong></td>
<td>79.3 ± 53.5</td>
<td>86.0 ± 50.2</td>
<td>1741</td>
<td>87.9 ± 44.6</td>
</tr>
</tbody>
</table>

*The group that manages includes pediatricians who have ever identified and referred and/or treated mothers for depression.
tricians who have not identified mothers with maternal depression, on average, pediatricians who have identified mothers with depression were significantly more likely to be older, to have been in practice for a longer period of time, and to be white. Pediatricians who have practices that assign patients to a specific pediatrician and those who see primarily white patients (≥75%) were also significantly more likely to identify mothers with depression. These 2 groups did not differ significantly with respect to pediatrician’s gender, location or region of practice, proportion of patients with private health insurance, or number of ambulatory visits per week.

Descriptive statistics for pediatricians who have identified and managed (referred and/or treated) maternal depression and pediatricians who have only identified maternal depression are also presented in Table 1. The results showed that the only pediatrician and practice characteristics associated significantly with the identification and management of maternal depression were the pediatrician’s race and region of practice; white pediatricians and pediatricians practicing in the Midwest were more likely to have managed maternal depression.

Table 2 shows the associations between pediatricians’ attitudes and practices toward mental health and their behaviors regarding mothers with depression for each outcome. Compared with pediatricians who have not identified maternal depression, those who have identified maternal depression were more likely to work in practices that provide child mental health services and to report using ≥1 method to address maternal depression. Pediatricians who perceived that adult mental health services are very or somewhat available, believed that maternal depression has an extreme or great effect on children’s physical health and mental health, or believed that caregiving problems attributable to maternal health have an extreme or great effect on children’s mental health were significantly more likely to identify maternal depression. Interestingly, perception of great or extreme effects of maternal caregiving problems on children’s physical health and mental health, or believed that caregiving problems attributable to maternal health were significantly more likely to identify maternal depression. Interestingly, perception of great or extreme effects of maternal caregiving problems on children’s physical health and mental health was not associated significantly with identification. The perception that pediatricians should be responsible for identification of but not referrals for or treatment of maternal depression was associated significantly with identification. Pediatricians who have identified maternal depression had significantly more-favorable attitudes regarding acknowledgment of maternal depression and perception of mothers’ beliefs, compared with pediatricians who have not identified maternal depression.

Similar associations were observed for comparisons of pediatricians who have only identified maternal depression and those who have identified and managed maternal depression (Table 2). However, there were some differences. Perceived availability of adult mental health services was not associated significantly with management of maternal depression. In addition, the perception of great or extreme effects of maternal caregiving problems on children’s physical health and mental health was associated significantly with management of maternal depression. Pediatricians who thought that pediatricians should be responsible for identification of maternal depression were significantly more likely to have managed maternal depression. Pediatricians who identified and managed maternal depression were more likely to agree that pediatricians should be responsible for making referrals, compared with pediatricians who only identified maternal depression; however, this did not reach statistical significance (P = .056). Of note, 77% of pediatricians who reported identifying but not managing maternal depression agreed that pediatricians should be responsible for making referrals for this condition. Finally, pediatricians’ mean attitude scores for perceptions of mothers’ beliefs and acknowledgment of maternal depression were significantly higher among pediatricians who identify and manage maternal depression, compared with those who only identify maternal depression. Pediatricians who identify and manage maternal depression tended to have higher attitude scores, compared with physicians who only identify maternal depression, but this difference was not statistically significant. Table 3 presents the frequency of use of methods to identify maternal depression, listed from least active (observation) to most active (screening instrument). Observation and family history were methods most often used by pediatricians. Screening instruments were rarely used; 13% of pediatricians who identify but do not manage maternal depression and 23% of pediatricians who manage maternal depression reported using a screening tool. Pediatricians who have managed maternal depression were significantly more likely to inquire about symptoms as part of routine discussion and to report that mothers inquire about depression or volunteer information about symptoms, compared with those who have only identified depressed mothers.

The results of the multivariate weighted logistic regression analyses modeling the odds of identifying a mother with maternal depression are shown in Table 4. Pediatrician and practice characteristics that were associated significantly with identification included pediatrician age and patient race. For each 5-year increase in the age of the pediatrician, the odds of identifying a mother with maternal depression increased by 27%. Pediatricians whose practice consisted of ≥75% white patients were twice as likely (OR: 2.07) to have identified a mother with maternal depression compared with pediatricians with <75% white patients. Several mental health attitudes and practices were also associated significantly with identifying maternal depression. Pediatricians who worked in practices that provide child mental health services (OR: 1.85), used ≥1 method to address maternal depression (OR: 2.55), agreed that pe-
Pediatricians should be responsible for identifying maternal depression (OR: 1.62), or thought that maternal depression has an extreme effect on children’s mental health (OR: 1.78) had significantly higher odds of identifying maternal depression. Moreover, more-favorable attitudes about the perception of mothers’ beliefs about pediatricians’ involvement in maternal depression were associated positively with identification of maternal depression (OR: 1.27). Two-way interactions of clinical relevance (eg, patient race and responsibility for identification and effect of children’s mental health) were evaluated and were not statistically significant ($P > .10$).

Table 5 reports the results of the multivariate, weighted, logistic regression analyses modeling the odds...
of identifying and managing maternal depression. Similar to the analyses of the outcome of identification, several mental health attitudes and methods were associated significantly with the outcome of management of maternal depression; pediatricians who worked in offices that provide child mental health services (OR: 2.71), pediatricians who used maternal depression; pediatricians who worked in offices that provide child mental health services (OR: 2.71), or who usually inquired about symptoms or contributing factors as part of routine discussion to identify maternal depression had twice the odds (OR: 2.09) of managing it. Because of the complexity and severity of maternal depression, intervention should occur early. Pediatricians who thought that caregiving problems attributable to maternal health have an extreme effect on children’s physical health were 3 times more likely to report managing maternal depression, compared with pediatricians who did not think that the effect is extreme. The results show that pediatricians who usually inquired about symptoms or contributing factors as part of routine discussion to identify maternal depression had twice the odds (OR: 2.09) of managing a mother with maternal depression. Unlike the identification outcome, the only physician or practice characteristic that was associated significantly with the management outcome was region of practice. Pediatricians practicing in the Midwest were significantly more likely to report managing maternal depression, compared with each of the other regions (compared with South: OR: 3.97; compared with West: OR: 5.16; compared with Northeast: OR: 5.06). Pediatricians’ attitudes about acknowledging maternal depression were also associated significantly with management of maternal depression.

DISCUSSION
In this representative sample of practicing pediatricians, more than three fourths reported identifying maternal depression, and 82% of those who do identify also manage mothers with depression (refer or treat). The overwhelming number of pediatricians who manage mothers with depression do so by referring them for additional assessment and treatment: very few (4%) reported that they provide treatment themselves. Pediatricians’ age and the racial distribution of their patients were associated with identification but not management of maternal depression. Practice characteristics (provision of mental health services and use of screening methods) were associated with identification and management of maternal depression. Pediatricians’ attitudes toward maternal depression, as measured by perceptions of mothers’ beliefs and acknowledgment of maternal depression, and their perceptions about the effect that maternal mental health has on children also affected pediatricians’ behaviors.

There was also some effect of experience. Older pediatricians and those who had been in practice longer were more likely to report identifying and managing mothers with depression. These covariates were highly correlated, and both might have been related to exposure to mothers with depression. Interestingly, neither was associated with management of mothers in logistic models controlling for other factors. The fact that age and years of experience in practice were not significant correlates of identification or management of maternal depression suggests that experience may increase a pediatrician’s facility with initiating discussion about maternal depression but the “next steps” of referral or treatment do not occur more often. It is important, therefore, to develop strategies that not only increase recognition of mothers with depression but also enable pediatricians to refer mothers at risk.

Pediatricians’ attitudes drive behavior for both identification and management of maternal depression. If pediatricians think that a problem is important for children’s well-being and they feel that they are responsible for it, then they act to identify and to manage it, as in the case of maternal depression. Previous research suggested that perceived barriers to mental health care, such as lack of time, lack of availability of mental health resources for referrals, and inadequate training to identify and to manage maternal depression may affect best prac-
Additional study is needed to disaggregate the interactions that influence pediatricians to identify and to manage mothers with depression.

It is not surprising to find that, compared with pediatricians who did not identify mothers with depression, respondents who thought that pediatricians should be responsible for identification were more likely to identify mothers with maternal depression. What is surprising is that the proportions of physicians who agreed that pediatricians should be responsible for making referrals for or treating maternal depression were similar regardless of whether they do not identify, only identify, or identify and manage maternal depression in practice. In fact, the majority of respondents thought that pediatricians should be responsible for referring mothers with depression. These findings are consistent with those of Olson et al, who reported that about 58% of pediatricians agreed that it is the responsibility of pediatricians to identify maternal depression, 83% of pediatricians agreed that it is the responsibility of pediatricians to refer patients for maternal depression, and only 5% of pediatricians thought that it is the responsibility of pediatricians to treat maternal depression. Therefore, it is important to ascertain the reason for lack of action on the part of pediatricians who think that it is their respon-

| TABLE 4 | Weighted Logistic Regression Model Results for Pediatricians Who Have Identified ≥1 Mother With Maternal Depression |
| --- | --- | --- |
| Covariates in the Model | OR (95% CI) | P |
| Pediatrician’s age (per 5-y increase) | 1.27 (1.13–1.44) | <.0001 |
| Patient ethnicity | | |
| <75% white (reference) | 1.00 | |
| ≥75% white | 2.07 (1.24–3.47) | .0057 |
| Methods used to address maternal depression | | |
| No methods (reference) | 1.00 | |
| ≥1 method | 2.55 (1.61–4.03) | <.0001 |
| Pediatrician or others in practice provide child mental health services | | |
| Do not provide services (reference) | 1.00 | |
| Provide services | 1.85 (1.15–2.99) | .0113 |
| Pediatricians should be responsible for identifying maternal depression | | |
| Neutral/disagree (reference) | 1.00 | |
| Agree | 1.62 (1.04–2.51) | .0326 |
| Effect of maternal depression on child’s mental health | | |
| Moderate/little/none/no answer (reference) | 1.00 | |
| Extreme | 1.78 (1.11–2.85) | .0168 |
| Perception of mothers’ beliefs on attitude scale | | |
| 1.27 (1.14–1.42) | <.0001 |

| TABLE 5 | Weighted Logistic Regression Results for Pediatricians Who Have Identified and Managed (Referred and/or Treated) Maternal Depression |
| --- | --- | --- |
| Covariates in the Model | OR (95% CI) | P |
| Region of practice | | |
| Northeast (reference) | 1.00 | |
| Midwest | 5.06 (1.69–15.16) | .0025 |
| South | 1.27 (0.65–2.51) | |
| West | 0.98 (0.48–2.02) | .0232 |
| Methods used to address maternal depression | | |
| No methods (reference) | 1.00 | |
| ≥1 method | 2.36 (1.35–4.11) | .0025 |
| Pediatrician or others in practice provide child mental health services | | |
| Do not provide services (reference) | 1.00 | |
| Provide services | 2.71 (1.52–4.82) | .0007 |
| Effect of caregiving problems attributable to maternal physical or mental health on child’s physical health | | |
| Moderate/little/none/no answer (reference) | 1.00 | |
| Extreme/great | 3.00 (1.42–6.34) | .0040 |
| Inquires about symptoms/contributing factors as part of routine discussion to identify maternal depression | | |
| Sometimes/never/no answer (reference) | 1.00 | |
| Usually | 2.09 (1.18–3.72) | .0119 |
| Acknowledging maternal depression on attitude scale | | |
| 1.28 (1.10–1.48) | .0010 |

a The Midwest was significantly different from each of the other regions (compared with South: OR: 3.97; 95% CI: 1.36–11.61; compared with West: OR: 5.16; 95% CI: 1.71–15.51). The remaining pairwise comparisons among the other 3 regions were not statistically significant.
sibility to identify and to treat maternal depression. This finding warrants careful investigation because, without identification, no appropriate referrals and treatment are possible.

In our sample, pediatricians in the Midwest were more likely to identify and to manage mothers with depression than were respondents from other geographic areas. Several highly populated Midwestern states (eg, Wisconsin, Michigan, and Illinois) have implemented large public-awareness campaigns about maternal depression.57–59 These statewide programs include not only information but also enhanced treatment options in the mental health community for maternal (postpartum) depression. As a result, pediatricians may have increased knowledge regarding postpartum depression and may have enhanced networks for referral of mothers. Additional research is needed to explain this interesting finding.

One of the interesting findings of this study is that pediatricians who use ≥1 method to address maternal depression were 2.36 times more likely to report having managed (referred or treated) an at-risk mother. Many pediatricians use family history as a means to identify women at risk, because family and personal histories of depression and postpartum depression are known risk factors and may be very helpful in identifying women at risk. The primary method reported for identifying maternal depression, that is, observation, has been shown repeatedly to be inadequate.40,41,60 Furthermore, more than one half of respondents do not inquire about symptoms, and more than three fourths do not use a screening tool. Both methods, when structured with specific depression questions, are more effective than observation or clinical assessment.61 In fact, many validated and easy-to-use screening tools exist62–64 and would be helpful in identifying mothers with depression. Olson et al61 found that, although paper-based screening for maternal depression yielded 4 times the number of positive screens, compared with interviewer-based questioning, both increased discussion of maternal depression and subsequent referral of mothers for additional care. There are many methods available to pediatricians to identify maternal depression, but using ≥1 is the first step in assisting mothers at risk.

Patient race was associated significantly with identification of maternal depression. This might have important implications. Smedley et al65 suggested that bias or stereotyping might contribute to different patient care and health outcomes. Patient attitudes also affect physician practices. Although some research has shown that black patients show more-favorable attitudes toward mental health treatment than do white patients,66 they are less likely to receive appropriate care, because of lack of insurance or distrust of health care providers.67 Other research has suggested that racial and cultural differences exist regarding the acceptability of antidepressant use or counseling for depression, with black subjects being less likely than white subjects to find antidepressant use acceptable and with Hispanic subjects being more inclined than white respondents to find counseling acceptable.68 The findings in our study suggest that the underlying racial and cultural differences among both patients and health care professionals in identifying and managing maternal depression warrant additional study.

The findings from this study must be evaluated in light of their limitations. First, this survey, like others of physicians, had a suboptimal (although solid) response rate.69,70 Although a detailed analysis of response rates in AAP surveys showed little nonresponse bias,71 the analyses were weighted for nonresponse; it is possible that this did not correct fully for nonresponse bias. Furthermore, pediatricians who are most interested in this topic are most likely to respond, and the results must be viewed with this fact in mind. Second, this was a cross-sectional, self-reported survey, rather than direct observation of pediatricians’ practices. Given the possibility of response biases for socially desirable behaviors, it may be that respondents overestimated some of the behaviors or actions listed on the survey. Additional study is warranted to test the associations found in this nationally representative sample of pediatricians, preferably with directly observed care72 or other prospective methods.

CONCLUSIONS
Pediatricians’ practice characteristics and attitudes are associated with their identification and management of mothers with depression. A significant shift in pediatricians’ attitudes toward mental health issues is necessary to improve recognition and care of depressed mothers. Training and education are essential, especially to encourage systematic screening for maternal depression, a condition that is amenable to treatment, which can reduce the duration of the depressive episode and the likelihood of adverse outcomes for both the child and the family. In fact, pediatricians are in an ideal position to screen for and to intervene in a myriad of conditions that affect mothers and their children, including children’s mental health problems, parental substance abuse, and domestic violence. Addressing such issues in the context of pediatric primary care may be uncomfortable for many pediatricians, but early systematic education, enhanced methods of identification that use technological advancements, and improved systems of care that ensure appropriate treatment for mothers and their children must be the goal for pediatricians in the 21st century.

ACKNOWLEDGMENTS
This study was supported by the AAP. This research was funded by the Annie E. Casey Foundation.
REFERENCES


43. Horwitz S, Kelleher K, Stein R, et al. Barriers to the identification and management of psychosocial issues in children and
maternal depression. *Pediatrics.* 2007;119(1). Available at: www.pediatrics.org/cgi/content/full/119/1/e208


72. Stange KC, Group TDW. The Direct Observation of Primary Care Study: insights from the process of conducting multi-method, transdisciplinary research in community family practice. *J Fam Pract.* 2001;50:345–352
The Influence of the Timing of Cord Clamping on Postnatal Cerebral Oxygenation in Preterm Neonates: A Randomized, Controlled Trial

Oskar Baenziger, MD, Florian Stolkin, MD, Mathias Keel, RN, Kurt von Siebenthal, MD, Jean-Claude Fauchere, MD, Seema Das Kundu, MD, Vera Dietz, MD, Hans-Ulrich Bucher, MD, Martin Wolf, PhD

Department of Neonatology, University Hospital Zurich, Zurich, Switzerland

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

ABSTRACT

OBJECTIVE. Our goal was to investigate the effect of placentofetal transfusion on cerebral oxygenation in preterm infants by near-infrared spectroscopy.

SUBJECTS. A total of 39 preterm infants with a median gestational age of 30.4 weeks were randomly assigned to an experiment group (n = 15) and a control group (n = 24).

INTERVENTIONS. The delivery of the infants in the experiment group was immediately followed by maternal administration of syntocinon, the infant was placed 15 cm below the placenta, and cord clamping was delayed by 60 to 90 seconds. The infants in the control group were delivered conventionally. At the ages of 4 and 24 hours, cerebral hemoglobin concentrations, cerebral blood volume, and regional tissue oxygenation were measured by near-infrared spectroscopy.

RESULTS. Cerebral blood volume was not different between the 2 groups at the age of 4 hours (6.1 vs 5.8 mL/100 g of tissue) nor at the age of 24 hours (6.2 vs 6.2 mL/100 g of tissue). Mean regional tissue oxygenation of the experiment group was higher at the ages of 4 hours (69.9% vs 65.5%) and of 24 hours (71.3% vs 68.1%).

CONCLUSION. Delayed clamping of the umbilical cord improves cerebral oxygenation in preterm infants in the first 24 hours.
In recent years, delayed cord clamping and placental fetal transfusion (PFT) was advocated to reduce the rate of blood transfusion and improve hemodynamic stability by increasing the intravascular blood volume in preterm neonates.\(^1\)\(^-\)\(^4\) PFT was proven to be a safe and practicable procedure and effectively can reduce the need for blood transfusion in very low birth weight infants.\(^5\)\(^,\)\(^6\) The increased blood volume may have profound implications on cerebral blood flow and oxygen delivery to the brain and other organs.\(^7\)\(^,\)\(^8\) This is particularly interesting because neonatal brain injuries are the most important sequelae of prematurity birth.

The aim of our study was to investigate the effect of PFT on cerebral oxygenation in preterm infants measured by near-infrared spectroscopy (NIRS).

**Patients and Methods**

A total of 39 preterm neonates expected to be delivered at 24 to 32 completed weeks’ gestation at the University Hospital Zurich were enrolled onto the study. The women were enrolled by a local study coordinator who obtained written informed consent and informed the obstetrician about the grouping. The neonatologist was not aware of the timing of cord clamping. Multiple deliveries, children with perinatal asphyxia, major fetal malformations, and children whose parents refused consent were excluded.

Our study population was a part of a large international, randomized, multicenter trial that investigated the effects of PFT on blood volume, need for red cell transfusion, and respiratory and neurologic complications. The infants were selected randomly and assigned to an experiment group or a control group by a central study coordinator. The infants were enrolled between September 1996 and July 1997.

The delivery of the infants in the experiment group was immediately followed by maternal administration of syntocinon, the infant was placed 15 cm below the placenta in cesarean-section deliveries and as low as possible for vaginal deliveries, and umbilical cord clamping was delayed 60 to 90 seconds. The infants in the control group were delivered conventionally, with the cord clamped in <20 seconds. The following neonatal resuscitation was performed according to the standard protocol of the clinic of neonatology and was identical for both groups. Clinical data about the 39 infants are provided in Table 1. There were no differences between the 2 groups in birth weight, gestational age, Apgar scores, or complications during pregnancy.

Eleven infants from the experiment group were delivered by cesarean section, and 16 infants from the control group were delivered by cesarean section. Three infants subsequently died (2 because of complications from hyaline membrane disease and 1 because of neonatal sepsis); all 3 were in the control group and died after the age of 72 hours.

The study was approved by the local ethics committee, and written informed consent was obtained from the parents before the study.

**Measurements**

At the ages of 4, 24, and 72 hours, cerebral oxygenation was evaluated by measuring deoxyhemoglobin (\(\mu\)M), oxyhemoglobin (\(\mu\)M), their sum of the total hemoglobin (tHb; \(\mu\)M), and the regional tissue oxygen saturation (StO\(_2\); %) by NIRS. From the tHb and the hematocrit, the cerebral blood volume (CBV; mL/100g) was calculated. The measurements were performed by trained study personal (Mr Keel and Drs Stolkin and Wolf) who participated in numerous NIRS studies, including studies on data quality and interobserver variability. At the same time heart rate, arterial blood pressure, hematocrit, arterial oxygen saturation, Pa\(_2\), and PaCO\(_2\) were recorded. For mechanically ventilated children, an oxygenation index was calculated by using the following formula:

\[
\text{oxygenation index} = \text{mean airway pressure} \times \frac{\text{fraction of inspired oxygen}}{100/\text{Pao}_2}
\]

**NIRS Instrument**

For our study, we used a Critikon 2020 Cerebral RedOx Monitor (Johnson & Johnson, New Brunswick, NJ), which is based on a 2-channel sensor and a coupling compensation system. It uses 4 laser diodes with wavelengths at 776.5, 819.0, 871.4, and 908.7 nm. The silicon photodiode detectors are placed 10 and 37 mm from the emitter window for the neonatal sensor. In the middle between the 2, a light emitting diode is placed. The light intensity of the light emitting diode should be the same at both detectors. Hence, differences in coupling can be compensated.

**NIRS Algorithm**

The goal of the Critikon algorithm is to determine the cerebral concentrations of hemoglobin without the influence of the superficial layers: skin, skull, and cerebrospinal fluid. The signal from detector 1 is mainly affected by the superficial layers, whereas the signal at detector 2

---

**Table 1 Clinical Data**

<table>
<thead>
<tr>
<th></th>
<th>Experiment Group (N = 15)</th>
<th>Control Group (N = 24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth weight, g</td>
<td>1115 ± 343.71</td>
<td>1330 ± 483.95</td>
</tr>
<tr>
<td>Gestational age, wk</td>
<td>30(\frac{3}{7}) ± 2.30</td>
<td>29(\frac{5}{7}) ± 2.38</td>
</tr>
<tr>
<td>Apgar, 1 min</td>
<td>5 ± 2.00</td>
<td>5 ± 2.67</td>
</tr>
<tr>
<td>Apgar, 5 min</td>
<td>8 ± 1.67</td>
<td>8 ± 1.77</td>
</tr>
<tr>
<td>Apgar, 10 min</td>
<td>8 ± 1.46</td>
<td>8 ± 1.87</td>
</tr>
<tr>
<td>Length at birth, cm</td>
<td>37.5 ± 4.23</td>
<td>39 ± 4.22</td>
</tr>
<tr>
<td>Head circumference, cm</td>
<td>27 ± 1.48</td>
<td>27 ± 2.77</td>
</tr>
</tbody>
</table>
has a substantial component that refers to brain. A ratio of the signals at detector 2 and detector 1 is calculated, which reduces the influence of the outer layers. The algorithm has been described in detail and yields absolute oxyhemoglobin, deoxyhemoglobin, tHb, and StO2 values. In previous studies, it was shown that the StO2 values obtained by this instrument reflect physiologic changes, although the instrument underestimates the size of the changes. The tHb has been compared with other methods of measurement.

**Statistics**

Differences between the groups were analyzed by using unpaired nonparametric tests (Mann-Whitney U); the results are expressed as mean and SD. The analyses were performed by using SPSS 7.5.1 (SPSS Inc, Chicago, IL).

**RESULTS**

Hematocrit was higher for infants in the experiment group compared with the control group at 4, 24, and 72 hours of age (Table 2). Mean arterial blood pressure was higher in the experiment group compared with the control group at 4 hours but did not differ at 24 and 72 hours (Table 2). Arterial oxygen saturation was lower in infants in the experiment group at the age of 4 hours but not in older infants. There was no difference between the experiment group and the control group regarding heart rate and PaCO2 at either age group.

Six infants from the experiment group and 12 infants from the control group needed mechanical ventilation. The oxygenation index between the 2 groups did not differ (13.63 ± 7.68 vs 16.0 ± 15.24).

**NIRS DATA**

A total of 15 infants in the experiment group and 24 infants in the control group were measured at the age of 4 hours. Because of movement artifacts or clinical instability of the child, 1 measurement at 24 hours (control group) and 3 measurements at age 72 hours (1 in the experiment group, 2 in the control group) could not be performed or the data set had to be excluded. At the age of 36 weeks, 12 measurements in the experiment group and 18 in the control group were performed (3 deaths in the control group or transfer to other neonatal units for the other infants).

The values of deoxyhemoglobin, oxyhemoglobin, CBV, and StO2 are given in Table 2. StO2 was higher at ages 4 and 24 hours in the experiment group but not at 72 hours (Table 3). There was a trend toward higher oxyhemoglobin values at the age of 4 hours (80.09 ± 34.91 vs 67.66 ± 29.32 μmol/L; P = .08), whereas there was no difference between the 2 groups in older infants. CBV and deoxyhemoglobin did not differ between the 2 groups at any age groups (Table 3).
DISCUSSION

The late clamping procedure led to a higher hematocrit measurement. Because the concentration of the oxygen carrier is higher, the oxygen extraction is lower, which led to the significantly higher StO₂ in the experiment group.

Delayed clamping of the umbilical cord allows PFT and improves oxygen delivery to the tissues by increasing systemic blood volume. Increased blood volume was advocated to facilitate pulmonary adaptation with a decrease for medical interventions, particularly mechanical ventilation.

PFT does increase total blood volume in preterm infants delivered by cesarean section. In infants in our experiment group, hematocrit and mean arterial blood pressure were increased compared with those in the control group, which suggests better systemic oxygen delivery to the tissue. It is well known that heart rate and arterial blood pressure only poorly correlate with blood volume, and that hematocrit does not completely match the adequacy of the oxygen transport capacity. However, our results demonstrate that the delayed cord clamping had its effect on systemic oxygen delivery to the tissues.

Optimizing tissue perfusion by PFT should also influence cerebral perfusion and potentially reduce the risk of hypoxic ischemic brain damage in these infants. We measured absolute cerebral hemoglobin concentrations by using NIRS and calculated a StO₂ that reflects tissue oxygenation of the brain. Our results demonstrate a higher StO₂ for the infants after delayed clamping of the umbilical cord. These results are supported by the data from studies using NIRS after red blood transfusion in older preterm infants. Transfusing red blood cells to anemic preterm infants resulted in an increase of oxyhemoglobin, deoxyhemoglobin, and the calculated StO₂. The main difference between our study and the above-mentioned study is that in the second study, adult red blood cells were transfused to anemic infants, whereas in our study, the transfusion was with the infant’s own fetal hemoglobin containing red blood cells.

The higher StO₂ measurement clearly demonstrates a higher cerebral tissue oxygenation in the experiment group. Even if the absolute values of StO₂ did not completely match the in vivo values of the tissue oxygenation, the difference between the 2 groups shows a greater reserve of cerebral oxygenation for the experiment group. This reserve potentially reduces the risk of hypoxic ischemic events to the brain.

Most authors suggest practicing PFT to reduce the need for transfusion of packed red cells. However, our data suggest that it may be beneficial to reduce the risk of disturbed cerebral oxygenation. Inadequate cerebral oxygenation is an important factor in the development of neonatal brain injury. In the preterm infant, it may play a role in the development of intracranial hemor-
rhages and periventricular leukomalacia. Our study demonstrates an increased tissue oxygenation in the neonatal brain after delayed cord clamping. We do not know whether this effect really has an impact on the development of neonatal brain injury or on the neurologic outcome of these infants, but reducing the risk of inadequate tissue oxygenation by a simple technique, such as delayed cord clamping, would be very advantageous.

Because the randomization for our study was performed by a central study coordinator for a larger randomized, multicenter trial and the primary outcome of the study was not cerebral oxygenation, the experiment group and the control group were not equal in size. They did not differ in their clinical data, such as birth weight, gestational age, head circumference, or Apgar scores. Therefore, the small study size and the uneven distribution should not affect the results in regards to cerebral oxygenation. However, it remains a limitation of our trial. A confirmation of the results by a larger, multicenter trial is desirable.

We conclude that delayed cord clamping increases cerebral oxygenation for the first 24 hours after birth.

REFERENCES

A Randomized, Controlled Trial of Acetaminophen, Ibuprofen, and Codeine for Acute Pain Relief in Children With Musculoskeletal Trauma

Eric Clark, MD, Amy C. Plint, MD, Rhonda Correll, BScN, Isabelle Gaboury, MSc, Brett Passi, MD

Departments of Pediatrics and Emergency Medicine, University of Ottawa, Ottawa, Ontario, Canada; Division of Emergency Medicine and Chalmers’ Research Group, Children’s Hospital of Eastern Ontario, Ottawa, Ontario, Canada; Faculty of Health Sciences, Queen’s University, Kingston, Ontario, Canada

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

ABSTRACT

OBJECTIVE. Our goal was to determine which of 3 analgesics, acetaminophen, ibuprofen, or codeine, given as a single dose, provides the most efficacious analgesia for children presenting to the emergency department with pain from acute musculoskeletal injuries.

PATIENTS AND METHODS. Children 6 to 17 years old with pain from a musculoskeletal injury (to extremities, neck, and back) that occurred in the preceding 48 hours before presentation in the emergency department were randomly assigned to receive orally 15 mg/kg acetaminophen, 10 mg/kg ibuprofen, or 1 mg/kg codeine. Children, parents, and the research assistants were blinded to group assignment. The primary outcome was change in pain from baseline to 60 minutes after treatment with study medication as measured by using a visual analog scale.

RESULTS. A total of 336 patients were randomly assigned, and 300 were included in the analysis of the primary outcome (100 in the acetaminophen group, 100 in the ibuprofen group, and 100 in the codeine group). Study groups were similar in age, gender, final diagnosis, previous analgesic given, and baseline pain score. Patients in the ibuprofen group had a significantly greater improvement in pain score (mean decrease: 24 mm) than those in the codeine (mean decrease: 11 mm) and acetaminophen (mean decrease: 12 mm) groups at 60 minutes. In addition, at 60 minutes more patients in the ibuprofen group achieved adequate analgesia (as defined by a visual analog scale <30 mm) than the other 2 groups. There was no significant difference between patients in the codeine and acetaminophen groups in the change in pain score at any time period or in the number of patients achieving adequate analgesia.

CONCLUSIONS. For the treatment of acute traumatic musculoskeletal injuries, ibuprofen provides the best analgesia among the 3 study medications.
Patients come to the emergency department (ED) with a variety of painful conditions, including fractures, bruises, and sprains. Within our pediatric ED, ~10% of all ED visits are for such injuries. Although providing adequate analgesia should be an important part of the ED treatment plan, numerous studies have shown that analgesia is not adequately provided to both pediatric and adult ED patients.1–6

When children are treated for pain in the ED, oral medications may be preferred. They eliminate the distress to a child of an intravenous or intramuscular injection and have a lower risk of the serious adverse events (such as apnea and aspiration) that are associated with parenteral pain medications. Although there have been studies comparing the pain relief provided by different oral analgesics in children postoperatively,7–10 there are no published randomized, controlled trials (RCTs) examining the use of common oral pain medications for children with acute musculoskeletal injury in the ED. Most published studies of oral analgesia for acute musculoskeletal pain in adult ED patients do not examine the oral analgesic agents commonly prescribed for children.11–14 One recent large ED-based study found no difference in pain relief between adult patients treated with paracetamol, indomethacin, diclofenac, or paracetamol combined with either nonsteroidal medication.15

The objective of this study was to determine which of 3 oral medications, acetaminophen, ibuprofen or codeine, given as a single dose, provides the most efficacious analgesia for children presenting to the ED with acute musculoskeletal traumatic injuries.

METHODS

Study Design
In this RCT we compared the change in pain among children with acute musculoskeletal pain treated with acetaminophen, ibuprofen, and codeine.

Study Setting and Population
This trial was performed between May 2002 and January 2003 at an academic, tertiary care children’s hospital in Ottawa, Canada, with an annual ED census of 55 000/year (the Children’s Hospital of Eastern Ontario). Children 6 to 17 years old were eligible if they presented to the ED with pain from a musculoskeletal injury (to extremities, neck, and back) occurring in the preceding 48 hours. Children were excluded if they had a contraindication to a study drug, required resuscitation, had an open fracture, had an intravenous line in place, had received 1 of the study drugs in the preceding 4 hours for acetaminophen and codeine or 6 hours for ibuprofen, or had a significant cognitive impairment. Written, informed consent was obtained. Our institutional research board approved this study.

Study Protocol
A research assistant recruited participants in the ED for 8 hours daily during the study period. Once consent was obtained, baseline data and study measurements were recorded. Participants were then assigned randomly to 1 of 3 groups. Participants received either 15 mg/kg of acetaminophen (maximum dose: 650 mg), 10 mg/kg of ibuprofen (maximum dose: 600 mg) or 1 mg/kg of codeine (maximum dose: 60 mg) by mouth. These doses were chosen because they have been used in other analgesia and antipyretic trials,11,16 the Compendium of Pharmaceuticals and Specialties, our national standard reference for medications lists these doses as standard, and our institution’s research pharmacist confirmed these doses, including maximum doses, as standard and recommended doses. The randomization sequence was computer generated with a block size of 9. Sealed opaque envelopes were used to conceal the allocation sequence. The drugs were all purple in color, grape flavored, and given in amber syringes covered with opaque plastic bags. Because of the pharmacokinetics of the drugs, the volumes of the study drug per kilogram were similar but not identical. To maintain blinding, the triage nurse opened the randomization envelope and administered the appropriate study medication. The triage nurse was not otherwise involved in the study or in additional care of the patient. The child, parent, and research assistant were blinded to group assignment.

The use of a visual analog scale (VAS) for measuring pain was explained by the research assistant to the children. The children recorded their baseline pain score by using a VAS before randomization and the assigned study drug being administered (“time 0”). Additional pain measurements were determined every 30 minutes for 120 minutes by using the VAS, and the child was not able to view previous scores to prevent carry over bias. All children were asked at 60 minutes and every 30 minutes afterward whether they required any additional analgesia. Additional pain medication was withheld for 60 minutes after administration of the study drug. Participants discharged before 120 minutes were given materials to complete the remaining scores at the appropriate times and stamped self-addressed envelopes. Parents were contacted by telephone 2 days after their visit to determine any adverse events and encourage mailing of the data forms.

All interventions including physical examinations, additional medications, radiographs, splints, casts, and reductions that occurred during the patient ED visit were prospectively documented, as was discharge diagnosis. Adverse effects in the ED were screened by using an open-ended question, “Is there anything bothering you other than your pain?” At the 2-day follow-up, adverse effects were screened for by specific and open-ended questions. Just before ED discharge, the children, parents, and research assistants were asked which medica-
tion they thought had been given. The final diagnosis and patient disposition was determined by the attending emergency physician. Diagnoses were then broadly grouped into fractures and soft tissue injuries.

**Outcome Measurements**

Baseline measurements included age, gender, pain score, and previous analgesic use. The primary outcome was change in patient’s self-reported pain from baseline at 60 minutes after receipt of the study medication. Pain was measured by using a VAS (a 100-mm hatched line anchored at 1 end with a label stating “no pain” and at the other end a label stating “worst pain”). VASs have been used extensively in analgesic trials and are valid for children ≥6 years of age. The clinically important change for a VAS is considered to range from 9 to 18 mm. We chose 60 minutes after administration as the timing of the primary outcome because drugs would all have been absorbed and efficacious by that point. The child’s report of pain rather than the parents’ or the research assistant’s was chosen because it has been shown that parents and health care workers are not accurate when assessing a child’s pain. Secondary outcomes included the change in VAS from baseline at 30, 90, and 120 minutes, requirement for additional analgesia, and the number of patients achieving a VAS < 30 mm (defined as “adequate analgesia”) at 60 and 120 minutes. This last outcome was chosen because a previous study suggested that a pain score < 30 mm indicates adequate pain relief.

**Sample Size**

Previous studies have indicated that the minimal clinically significant difference in pain, as measured by a VAS, ranges from 9 to 18 mm with an SD ranging from 14 to 40 mm. Given this range, we chose a 15-mm difference (SD: 20 mm) in the change in VAS score between groups because of our minimal clinically significant difference to detect. Sample-size calculations were thus based on the following assumptions: (1) detection of a 15-mm difference between groups, (2) standard deviation of 20 mm, (3) 2-sided test, and (4) statistical power of 80% and false-positive (type I error) rate of 0.05.

Although these assumptions were appropriate, the formulae used to calculate the sample size was a posteriori found to be inadequate. First, a formula for a 2-arm trial was used and expanded to accommodate a 3-arm trial. Second, the sample size obtained by using the above assumptions required a total number of 56 participants, which was mistakenly interpreted as 56 participants per arm. Thus, we planned to enroll 168 patients in total and doubled that number to have sufficient power for the subgroup analyses.

**Data Analysis**

Gender and type of injury of enrolled versus nonenrolled eligible children were compared by using χ² tests. Difference in age was assessed by using Student’s t test. Comparison of continuous outcomes (such as change in VAS from baseline) between the 3 study groups was determined by using analysis of variance models, followed by Tukey post hoc tests of significance when a significant difference was observed. Categorical outcomes (such as adequate analgesia achieved) were compared by using χ² tests or Fisher’s exact tests when necessary. Success in blinding was assessed by using a χ² test. All reported P values were 2-sided and deemed significant when they reached a 5% level. A priori planned subgroups included those with baseline VAS measurements of > 30 mm (because they were assumed to have more “significant pain”), patients with fractures, and patients with soft tissue injuries.

Data were first analyzed on a per protocol basis. Patients were included in per protocol analysis if they received a dose of the study drug, had baseline data, and had primary outcome data. An intention-to-treat analysis, which included all patients initially randomly selected, was performed on the primary outcome and change in pain score from baseline at 120 minutes. Data for participants on whom a complete set of information was not available were imputed by using the last value carried forward.

**RESULTS**

**Patient Recruitment and Baseline Characteristics**

A total of 801 children with pain secondary to acute musculoskeletal injury presented to the ED during the time research assistants were available. Seven hundred eighty children were eligible, and 336 were enrolled (Fig 1). Three hundred twenty-four families refused to consent, 48 children were not approached because the research assistant was enrolling another child, 38 children were missed, and 34 were not enrolled for other reasons. Enrolled children were discharged from the ED throughout the study period, and the number for whom outcome data were available is indicated in Fig 1. Three hundred patients had a primary outcome measurement obtained for the final analysis (Fig 1). Enrolled versus nonenrolled eligible patients were comparable in age and gender, although not randomly selected patients were slightly more likely to have soft tissue injuries than randomly selected patients (54% vs 47%). Baseline characteristics were similar in all study groups (Table 1). Twenty-three patients (22%) in the ibuprofen group, 51 (48%) in the acetaminophen group, and 23 (21%) of children in the codeine group received the maximal study drug dosages based on weight.
Overall, patients showed improvement in pain from baseline over the course of the study. At 30 minutes, however, there was no significant difference in change in pain score among the 3 groups. From 60 minutes and onward, patients in the ibuprofen group had significantly greater improvement in pain score than those in the codeine and acetaminophen groups. There was no significant difference in the change in pain score between codeine and acetaminophen groups at any time period (Table 2). In addition, at 60 minutes more patients in the ibuprofen achieved adequate analgesia (as defined by a VAS <30 mm) than the other 2 groups. There was no statistical difference between the codeine and acetaminophen groups (Table 2). Over the course of the trial, there was no significant difference in the number of patients requiring additional analgesic (22.2% of codeine, 15.6% of acetaminophen, and 14.3% of ibuprofen).
Patients with soft tissue injuries

Patients with fractures

Subgroup Comparisons

Details of subgroup comparisons are reported in Table 3. Among patients with fractures, ibuprofen resulted in significantly better improvement in pain than the other medications at both 60 and 120 minutes. There was no statistical difference between codeine and acetaminophen. Among patients with a soft tissue injury, there was no significant difference in change in pain score among any of the 3 medications at 60 or 120 minutes. When only patients with pain >30 mm were considered in the analysis, ibuprofen was significantly better than the other medications at 60 minutes. The other drugs were equivalent. At 120 minutes, both ibuprofen and codeine had similar effects and were significantly better than acetaminophen.

Blinding

Patients and parents seemed to be adequately blinded to the identity of the study medication, choosing the correct response no greater than chance would allow. The research assistants, however, correctly identified the study drug as acetaminophen in 52% of cases and ibuprofen patients; $P = .32$). All of these medications were given after measurement of the primary outcome, thus the analysis was not adjusted for these additional treatments. The intention-to-treat analysis for change in pain score from baseline at 60 and 120 minutes and number of patients achieving adequate analgesic gave similar results to the per-protocol analysis (data not shown).

Adverse Effects and Adverse Events

No significant adverse effects were reported while study participants were in the ED. One child in the codeine group was accidentally administered 5 mg/kg of codeine as a single dose. This child was withdrawn from the study, treated with oral charcoal, monitored in the ED, and had no adverse outcome. At 48-hour telephone follow-up, there was no significant difference in the number of patients reporting minor adverse effects (such as nausea, sleepiness, and constipation), with 16 (16.2%) of 99 patients in the codeine group, 8 (7.7%) of 104 patients in the acetaminophen group, and 11 (10.9%) of 101 patients in the ibuprofen group reporting ≥1 adverse effect ($P = .16$).

---

**TABLE 2** Change in Pain Score (VAS) From Baseline and Number of Patients Who Achieved Adequate Analgesia

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Codeine</th>
<th>Acetaminophen</th>
<th>Ibuprofen</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Mean or n (%)</td>
<td>95% CL</td>
<td>N</td>
</tr>
<tr>
<td>Change in VAS from baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 min</td>
<td>105</td>
<td>−10</td>
<td>−14, −6</td>
</tr>
<tr>
<td>60 min</td>
<td>100</td>
<td>−11</td>
<td>−16, −5</td>
</tr>
<tr>
<td>90 min</td>
<td>85</td>
<td>−13</td>
<td>−20, −6</td>
</tr>
<tr>
<td>120 min</td>
<td>75</td>
<td>−17</td>
<td>−25, −9</td>
</tr>
<tr>
<td>VAS &lt;30 mm&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60 min</td>
<td>100</td>
<td>40 (40)</td>
<td>31, 50</td>
</tr>
<tr>
<td>120 min</td>
<td>75</td>
<td>39 (52)</td>
<td>41, 63</td>
</tr>
</tbody>
</table>

Post hoc test significance (Tukey): VAS at 60 minutes: acetaminophen versus codeine ($P = .98$); acetaminophen versus ibuprofen ($P = .001$); codeine versus ibuprofen ($P < .001$); VAS at 90 minutes: acetaminophen versus codeine ($P = .35$); acetaminophen versus ibuprofen ($P = .016$); codeine versus ibuprofen ($P = .001$); VAS at 120 minutes: acetaminophen versus codeine ($P = .85$); acetaminophen versus ibuprofen ($P = .028$); codeine versus ibuprofen ($P = .006$); VAS <30 mm at 60 minutes: acetaminophen versus codeine ($P = .85$); acetaminophen versus ibuprofen ($P = .026$); codeine versus ibuprofen ($P = .006$). CL indicates confidence limit. Adequate analgesia was defined as a VAS <30 mm.

<sup>a</sup> The number of patients at each time decreased because patients were discharged home from the ED and did not return pain scores for later time periods.

<sup>b</sup> Because we have no data on patients who did not complete the VAS at home after discharge, these data represent cross-sectional data on those patients on whom VAS scores were available.

---

**TABLE 3** Mean Change in VAS From Baseline Among Patients With Fractures, Soft Tissue Injuries, and Baseline VAS >30 mm

| Subgroup<sup>a</sup> | Codeine | | Acetaminophen | | Ibuprofen | |
|----------------------|---------|-----------------|----------|-----------------|----------|
| n | Mean (95% CLs) | n | Mean (95% CLs) | n | Mean (95% CLs) |
| | | | | | |
| Patients with fractures | | | | | |
| 60 min | 50 | −7 (−8, −6) | 51 | −14 (−19, −9) | 58 | −29 (−35, −22) |
| 120 min | 42 | −13 (−24, −3) | 42 | −20 (−28, −13) | 48 | −41 (−49, −33) |
| Patients with soft tissue injuries | | | | | |
| 60 min | 50 | −14 (−22, −7) | 49 | −9 (−16, −2) | 42 | −19 (−24, −13) |
| 120 min | 33 | −22 (−34, −10) | 37 | −19 (−28, −9) | 35 | −18 (−26, −11) |
| Patients with VAS >30 mm at baseline | | | | | |
| 60 min | 74 | −18 (−24, −12) | 84 | −13 (−18, −8) | 89 | −27 (−32, −22) |
| 120 min | 54 | −27 (−35, −19) | 66 | −23 (−29, −17) | 77 | −34 (−40, −28) |

CL indicates confidence limit.

<sup>a</sup> The number of patients at each time decreased as patients were discharged home from the ED and did not return pain scores for later time periods.
proven in 42% of cases, which is greater than would be expected by chance.

**DISCUSSION**

ED visits for painful conditions such as fractures, bruises, and sprains are extremely common. Despite this, analgesia is often not adequately provided to patients and in particular, to pediatric patients.1–6 Our study, the first to our knowledge to compare 3 commonly used oral medications in the treatment of pain from musculoskeletal injuries in children, has shown that ibuprofen provides better acute pain relief for these children than acetaminophen or codeine. Children receiving ibuprofen were also more likely to obtain adequate analgesia. We found no difference in the number of adverse effects among the 3 medications. Interestingly, in the subgroup analysis ibuprofen resulted in a greater improvement in pain scores among patients with fracture compared with codeine or acetaminophen, but no statistical difference between the 3 medications was seen among patients with soft tissue injuries. Given that the baseline pain score among patients with soft tissue injuries was the same as that among patients with fractures, this cannot be related to a lower “room for improvement” among patients with soft tissue injuries. The etiology of this difference is not clear but may reflect a difference in physiology of pain between the 2 groups. Perhaps the antiinflammatory effects of ibuprofen are responsible for the better pain relief.

It is also important to note that although ibuprofen was more efficacious in providing adequate analgesia, only 52% of children in this group could be defined as receiving “adequate analgesia” at 60 minutes. In addition, although codeine and acetaminophen did result in some improvement in pain, the actual level of improvement (a change of 10–11 mm on the VAS) is only just within the range previous studies have suggested to represent a significant improvement in pain.18–22 Thus, although ibuprofen provided better pain relief than codeine and acetaminophen in our study, it seems that ibuprofen alone is not adequate for relieving pain in all children with musculoskeletal injuries.

Although no study has performed a direct comparison of these medications among children with musculoskeletal injuries, other studies have shown ibuprofen to be better than acetaminophen for other painful conditions.16,26 For example, ibuprofen has been shown to be superior to acetaminophen for pain control in tonsillitis16 and for pain related to migraines.16 Two studies have compared pain relief with ibuprofen to an acetaminophen-codeine combined preparation7,8 and post tonsillectomy. Results are conflicting, with 1 study suggesting similar pain relief6 and another suggesting that the combined acetaminophen-codeine preparation may be slightly better.7 Neither of these studies used ibuprofen at 10 mg/kg per dose. In a study of patients with acute low back pain, another nonsteroidal antiinflammatory drug (NSAID), oral ketorolac, was found to give no better pain relief than an acetaminophen-codeine preparation.11 In contrast to our results, many of these studies suggested the narcotic analgesics were associated with greater adverse effects than NSAIDs.8,10,11 Most of these studies, however, treated patients with multiple medication doses. A large (n = 300) ED study of adult patients with pain from acute musculoskeletal injury found no difference in pain relief among patients treated with paracetamol, indomethacin, diclofenac, or a combination of paracetamol and NSAIDs. The dose of indomethacin and diclofenac, although dosages commonly used, were not the maximum doses allowed. In addition, unlike our study, only a small number of patients had fractures.15 One small, nonrandomized 3-arm trial (76 patients) compared the effect of “standard care” (ice and elevation), “standard care” plus 10 mg/kg ibuprofen, and “standard care” plus distraction on pain relief in children with fractures. Interestingly, this trial found that ibuprofen added no pain relief benefit to standard care, although distraction was beneficial.27

There have been concerns expressed regarding the effect of NSAIDS on bone metabolism and fracture healing. Animal studies have suggested that multiple doses of indomethacin, aspirin, and ibuprofen27–30 can all affect the healing of variety of fractures in rats. Retrospective studies in humans have given inconsistent results. No prospective RCTs have examined the effect of ibuprofen on fracture healing. One RCT examining the use of piroxicam found no significant delay in healing of Colle’s fractures31 whereas another RCT found that a 6-week course of indomethacin significantly increased the risk of nonunion of acetabular fractures.32 There is no evidence that a single does of ibuprofen is associated with delayed fracture healing in humans. In addition, because NSAIDs inhibit platelet aggregation and prolong bleeding time, their use could increase the risk of bleeding. However, a recent systematic review found no increase in bleeding when NSAIDs were used for pain control post tonsillectomy.33 However, over-the-counter NSAIDs such as ibuprofen and naproxen have been associated with an increased risk of serious gastrointestinal toxicity (including gastrointestinal bleeds),34 although this risk seems to be related to length of usage.35 Limitations of this study include the large number of eligible patients who were not recruited for the study. The study patients were, however, similar to the non-enrolled patients with regards to their baseline characteristics, including age and gender, although they more likely to have fractures as their final diagnosis. Interestingly, the most common reason for refusal of consent was that the parents felt the child’s pain was not severe enough to justify pain medication. This suggests that education of parents regarding the benefits and efficacy of analgesics for children may be necessary. Although 36
randomly selected patients did not have primary outcome data, we had sufficient sample size to demonstrate a difference between study medications. Furthermore, although no difference was noted in adverse effects among the study groups, this study was not powered to detect rare, serious adverse events. In addition, the number of adverse effects reported may increase when a checklist is used for screening. In the ED, we used an open-ended question, although our 2-day follow-up included a checklist and open-ended question. More children in the acetaminophen group received the maximum study drug dose than in the ibuprofen and codeine group. Although we chose our maximum drugs doses on the basis of previous studies and standard doses, it is possible that the use of higher maximum doses of codeine or acetaminophen might have resulted in better pain relief with these medications.

This study may be further limited by the difficulty in blinding. Although the Consolidated Standards of Reporting Trials (CONSORT) statement recommends reporting “how the success of blinding was evaluated,” recently there has been debate regarding the correct way to assess the adequacy of blinding in RCTs. In our study, we asked patients, parents, and research assistants to guess which study medication was received. We found that the research assistant was correctly identified the study drug as acetaminophen in 52% of cases and ibuprofen in 42% of cases, which suggested that blinding may not have been adequate. However, we feel this does not invalidate the results in that neither the participants nor the parents seemed able to determine which study drug the child received, and the primary outcome was the child’s self-reported change in pain.

In conclusion, our study demonstrates that among children with pain from acute musculoskeletal injuries presenting to a pediatric ED, a single dose of ibuprofen provides greater pain relief than codeine or acetaminophen.

ACKNOWLEDGMENTS

This study was supported by a research grant from the Children’s Hospital of Eastern Ontario Research Institute. Dr Plint was supported in part by a salary-support award from Children’s Hospital of Eastern Ontario Research Institute.

REFERENCES

18. Kelly AM. Does the clinically significant difference in visual analog scale pain scores vary with gender, age or cause of pain? Acad Emerg Med. 1998;5:1086–1090
24. Singer AJ, Gulla J, Thode HC. Parents and practitioners are poor


39. Sackett DL. Turning a blind eye: why we don’t test for blindness at the end of our trials. *BMJ.* 2004;328:1136


PFIZER WILL REDUCE SALES FORCE

“Pfizer, the world’s largest drug company, said . . . that it would lay off almost 2400 sales representatives and managers, which is a fifth of its United States sales force. . . . The move may indicate the beginning of a wider retrenchment by Pfizer and the rest of the drug industry. Drug makers have sharply increased the size of their sales forces over the last decade as the research productivity of the companies has plunged and the pipeline of important new drugs has dwindled. The bloated sales forces, analysts say, have alienated doctors and contributed to high drug prices. Because Pfizer led the sales force expansion, other companies will probably follow its decision to cut back.”


Noted by JFL, MD
ABSTRACT

OBJECTIVE. The aim of our study was to examine long-term effects of nephrocalcinosis in prematurely born children.

PATIENTS AND METHODS. Preterm neonates (gestational age < 32 weeks) with \((n = 42)\) and without \((n = 32)\) nephrocalcinosis were prospectively studied at a mean age of 7.5 \((\pm 1.0)\) years.

RESULTS. Blood pressure did not differ in ex-preterm infants with and without nephrocalcinosis but was significantly higher than expected for healthy children. In comparison to healthy children, more ex-preterm infants with neonatal nephrocalcinosis had (mild) chronic renal insufficiency (glomerular filtration rate: <85 mL/min per 1.73 m\(^2\); 6 of 40); this is in contrast to ex-preterm infants without neonatal nephrocalcinosis (2 of 32). Tubular phosphate reabsorption and plasma bicarbonate were significantly lower in children with nephrocalcinosis compared with children without nephrocalcinosis. In addition, more ex-preterm infants with and without nephrocalcinosis than expected had low values for plasma bicarbonate and early-morning urine osmolality compared with healthy children. Kidney length of ex-preterm infants with and without nephrocalcinosis was significantly smaller than expected in healthy children of the same height. Nephrocalcinosis persisted long-term in 4 of 42 children but was not related to blood pressure, kidney length, or renal function.

CONCLUSIONS. Nephrocalcinosis in preterm neonates can have long-term sequelae for glomerular and tubular function. Furthermore, prematurity per se is associated with high blood pressure, relatively small kidneys, and (distal) tubular dysfunction. Long-term follow-up of blood pressure and renal glomerular and tubular function of preterm neonates, especially with neonatal nephrocalcinosis, seems warranted.
Nephrocalcinosis, first described in preterm neonates in 1982 by Hufnagel, was initially attributed to long-term furosemide therapy for chronic lung disease. Since then, nephrocalcinosis was diagnosed in 7% to 64% of preterm neonates with gestational age <32 weeks or birth weight <1500 g. The wide range in prevalence of nephrocalcinosis is a consequence of different study populations and ultrasound equipment and criteria, in addition to a moderate interobserver variation. Nephrocalcinosis in preterm neonates occurs as a result of imbalance between stone-promoting and stone-inhibiting factors. The etiology of nephrocalcinosis is multifactorial, consisting of low gestational age and birth weight, severe respiratory disease, high intakes of calcium, phosphate, and ascorbic acid; long duration of total parenteral nutrition; as well as drugs with hypercalciuric adverse effects, eg, furosemide, corticosteroids, and methylxanthines. Renal histology from autopsies of intensively treated neonates showed predominantly intratubular calciumoxalate and, in a minority of children, calcium phosphate deposits.

Resolution of nephrocalcinosis commonly occurs in the first years of life. In a previous study, we found nephrocalcinosis persisted in only 15% of patients beyond 30 months of age. However, the question arises whether the presence of nephrocalcinosis in a preterm neonate has unfavorable long-term effects on blood pressure and renal function. Data on follow-up of renal function of preterms with nephrocalcinosis are scarce, involves small numbers of patients, and results are not always compared with a control group of preterms without nephrocalcinosis. Here, we present the largest prospective long-term follow-up study up to date of blood pressure and renal glomerular and tubular function in preterm neonates with and without neonatal nephrocalcinosis.

PATIENTS AND METHODS

General Outline

In our previous prospective observational study, 41% of 201 preterm neonates with gestational age <32 weeks developed nephrocalcinosis at term. The aim of the current study was to analyze long-term blood pressure and renal function and compare results of ex-preterm infants with and without neonatal nephrocalcinosis. In addition, children were checked for persistence of nephrocalcinosis and excretion of calcium, and citrate was measured. Patients were born between May 1996 and November 1998 at Leiden University Medical Center and Haga Teaching Hospital, Juliana Children’s Hospital, Netherlands. All children alive with (n = 80) and without (n = 92) nephrocalcinosis at term were asked to participate in the study. Three children died, and children classified as having dubious nephrocalcinosis at term, defined as slight changes on renal ultrasound possibly but not unequivocally fitting nephrocalcinosis (n = 26), were excluded from the follow-up study. The ethics committee of both hospitals approved the study protocol. Informed consent was obtained after written information had been given.

Renal Function

Glomerular Function

Estimated glomerular filtration rate (GFR) was used to study renal function with a modified Schwartz formula (40 × length [cm]/serum creatinine [μmol/L]),. Serum creatinine concentration was measured by means of a photometric method (Jaffe) on an automatic analyzer (Modular P 800, Hitachi, Tokyo, Japan at Leiden University Medical Center and Synchroph L20-pro, Beckman Coulter, Fullerton, CA, at Juliana Children’s Hospital). Chronic renal insufficiency was defined as a GFR <85 mL/min per 1.73 m2. Urinary microalbumin was measured with an immunoturbidimetric assay on a fully automated Hitachi 911, and the variation coefficient ranged from 1.5% to 3.1% at different levels. The reference value for urine microalbumin/creatinine is <2.5 μg/μmol.

Tubular Function

Proximal tubular function was studied by means of tubular reabsorption of phosphate (TRP), plasma bicarbonate, and presence of glucosuria. Serum and urine phosphate concentration was measured on an automatic analyzer (Hitachi 747–100). Plasma bicarbonate was determined with a colorimetric assay on a fully automated Hitachi 911, with a variation coefficient of <2%. Urine anion gap (reference value: less than −20 to −50 mmol/L for adequate distal hydrogen ion production during metabolic acidosis) was used to differentiate between proximal and distal tubular acidosis. Urine glucose was tested semiquantitatively with a dipstick; a positive result corresponding to glucose was >0.3 g/L.

Distal tubular function was evaluated by means of early-morning urine osmolality and plasma bicarbonate in combination with urine anion gap. Urine osmolality was determined by testing freezing point depression.

Renal Ultrasound

Two pediatric radiologists, 1 in each center, performed ultrasound examinations with curved array broadband 5- to 12-MHz transducers, with additional imaging of details by means of small part transducers if possible.
Nephrocalcinosis was defined as the presence of bright reflections in the medulla or cortex, reproducible in both transverse and longitudinal direction with or without acoustic shadowing. Small bright reflections (≤2 mm) on renal ultrasound were classified as focal nephrocalcinosis and larger bright reflections (>2 mm) were classified as extensive nephrocalcinosis.

Intraobserver and interobserver agreement of ultrasound for the detection of nephrocalcinosis in preterm neonates by the 2 pediatric radiologists were previously found to be good (κ value: 0.84) and moderate (κ value: 0.46), respectively.9 Kidney length was compared with values for healthy children of the same height.29

Miscellaneous
Blood pressure was measured 3 times using an oscillometric device (Critikon Dynamap, GmbH, München, Germany). The lowest value was used and compared with values for healthy children of the same gender, age, and height.10 Urine analysis was performed by dipstick; if the dipstick was positive for hemoglobin or leukocytes, microscopy was performed. Urine calcium, creatinine, and citrate were measured on an automatic analyzer (Hitachi 747-100).

Statistics
Statistical significance of the variables was tested using χ² at a 5% significance level. χ² tests were also used to assess the fit of theoretical distributions specified by population percentiles. Group differences in baseline variables were tested by using χ² tests and independent samples t tests. All computations were conducted in SPSS 12.0 (SPSS Inc, Chicago, IL).

RESULTS
Patients
A total of 42 (53%) of 80 patients with nephrocalcinosis and 32 (35%) of 92 patients without nephrocalcinosis at term participated in the study. One patient without nephrocalcinosis at term is excluded because of cystic kidney disease. The remaining patients refused to participate in the study or were lost to follow-up. Patient characteristics are depicted in Table 1. There was no significant difference in gestational age or in birth weight between children with nephrocalcinosis at term participating (Table 1) and not participating (28.7 ± 1.9 weeks and 1136 ± 302 g; n = 38; P = .35) in the follow-up study. Likewise, there was no significant difference in gestational age or in birth weight between children without nephrocalcinosis at term participating (Table 1) and not participating (29.3 ± 1.7 weeks and 1264 ± 313 g; n = 58; P = .49) in the follow-up study.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Patient Characteristics, Blood Pressure, and Renal Function of Ex-preterm Infants With (NC⁺) and Without (NC⁻) Neonatal Nephrocalcinosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient characteristics</td>
<td>NC⁺</td>
</tr>
<tr>
<td>No. of patients</td>
<td>42</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>18 (43)</td>
</tr>
<tr>
<td>Gestational age, mean ± SD, wk</td>
<td>28.9 ± 2.3</td>
</tr>
<tr>
<td>Birth weight, mean ± SD, g</td>
<td>1148 ± 394</td>
</tr>
<tr>
<td>Birth weight, mean ± SD, SDS</td>
<td>-0.24 ± 1.10</td>
</tr>
<tr>
<td>Follow-up</td>
<td></td>
</tr>
<tr>
<td>Age, mean ± SD, y</td>
<td>7.4 ± 1.0</td>
</tr>
<tr>
<td>Weight for height, mean ± SD, SDS</td>
<td>-0.53 ± 1.48</td>
</tr>
<tr>
<td>Height, mean ± SD, SDS</td>
<td>-0.97 ± 1.21</td>
</tr>
<tr>
<td>Blood pressure, n (%)b</td>
<td></td>
</tr>
<tr>
<td>Systole</td>
<td></td>
</tr>
<tr>
<td>&gt;95th percentile</td>
<td>3 (7)</td>
</tr>
<tr>
<td>90th–95th percentile</td>
<td>7 (17)</td>
</tr>
<tr>
<td>50th–90th percentile</td>
<td>24 (57)</td>
</tr>
<tr>
<td>Diastole</td>
<td></td>
</tr>
<tr>
<td>&gt; 95th percentile</td>
<td>0 (0)</td>
</tr>
<tr>
<td>90th–95th percentile</td>
<td>1 (2)</td>
</tr>
<tr>
<td>50th–90th percentile</td>
<td>23 (55)</td>
</tr>
<tr>
<td>Nephrocalcinosis, n (%)</td>
<td>4 (9.5)</td>
</tr>
<tr>
<td>Glomerular filtration rate, mean ± SD, mL/min per 1.73 m²</td>
<td>108 ± 22</td>
</tr>
<tr>
<td>Chronic renal insufficiency, n (%)</td>
<td>6 (15)</td>
</tr>
<tr>
<td>Microalbuminuria, n (%)</td>
<td>3 (8)</td>
</tr>
</tbody>
</table>

a P < .05.
b Compared with healthy children, significantly more children had high blood pressure in both groups (P < .0001).
c Significantly more ex-preterm infants with nephrocalcinosis had chronic renal insufficiency (GFR < 85 mL/min per 1.73 m²) than expected in healthy children (P < .0001) in contrast to ex-preterm infants without nephrocalcinosis (P = .18).
Growth and Blood Pressure
Weight for height (SDS) and height (SDS) of the patients are depicted in Table 1. Systolic and diastolic blood pressure did not differ significantly in both groups but was significantly higher than expected for healthy children (Table 1). None of the children received antihypertensive medication at follow-up. Birth weight SDS was not related to high blood pressure (systole: $P = .13$; diastole: $P = .41$).

Glomerular Function
In comparison to healthy children, more ex-preterm infants with neonatal nephrocalcinosis had (mild) chronic renal insufficiency (GFR $<85 \text{ mL/min per 1.73 m}^2$) than expected (6 of 40; $P < .0001$); this is in contrast to ex-preterm infants without neonatal nephrocalcinosis (2 of 32; $P = .18$). However, mean GFR was not significantly different in ex-preterm infants with nephrocalcinosis compared with infants without neonatal nephrocalcinosis (Table 1; Fig 1). Patients with (mild) chronic renal insufficiency did not have relevant concomitant medical problems. The degree of nephrocalcinosis at term (18 of 42 had focal nephrocalcinosis, and 24 of 42 had extensive nephrocalcinosis) did not predict glomerular function at follow-up ($P = .62$).

Urine albumin excretion did not differ significantly between both groups. Furthermore, values in either group were within the reference range compared with values for healthy children from literature (Table 1).

Tubular Function
Proximal and distal tubular function are shown in Table 2. Plasma bicarbonate is classified under distal tubular function because the urine anion gap of all patients with low plasma bicarbonate was more than $-20 \text{ mmol/L}$, indicating inadequate distal hydrogen ion production during metabolic acidosis. Tubular function of children with extensive nephrocalcinosis at term (24 of 42) did not differ significantly from children with focal nephrocalcinosis (18 of 42).

Urine Analysis
Urinary excretion of calcium and citrate are depicted in Table 2. No patient had (microscopic) hematuria.

Renal Ultrasound
Kidney length of ex-preterm infants with and without neonatal nephrocalcinosis is depicted in Table 3. Kidneys of patients with as well as without neonatal nephrocalcinosis were significantly smaller than expected for healthy children, with the exception of left kidneys of children without neonatal nephrocalcinosis, which showed only a trend in being smaller ($P = .054$). Nephrocalcinosis persisted long-term in 4 (10%) of 42 children. Persistence of nephrocalcinosis was not related to blood pressure or kidney length. Only 1 child with persisting nephrocalcinosis had (mild) chronic renal insufficiency (GFR: 68 mL/min per 1.73 m$^2$). No child had low TRP or plasma bicarbonate, and only 1 had slightly decreased early-morning urine osmolality (675 mOsm/kg water). All 4 children with persisting nephrocalcinosis at long-term follow-up had a urinary calcium/creatinine ratio within the reference range, and only 1 child had a low urine citrate/creatinine ratio (0.07 mmol/mmol).

One child with neonatal nephrocalcinosis passed a few small kidney stones until the age of 3 years. At long-term follow-up, he had persistent nephrocalcinosis but no kidney stones.

DISCUSSION
General
With improvement of medical science and technical possibilities in the care of very prematurely born neonates,
quality of life of surviving children becomes increasingly important. There is growing evidence of an association of low birth weight with low nephron numbers and subsequent risk for adult cardiovascular disease and renal insufficiency. Development of nephrocalcinosis in prematurely born children may carry an additional risk of further compromising renal function later in life.

In our previous prospective observational study, 41% of 201 preterm neonates with gestational age of <32 weeks developed nephrocalcinosis at term. Here, we prospectively studied blood pressure, renal function, and long-term persistence of nephrocalcinosis in the same cohort of children and compared results with a control group of ex-preterm infants without nephrocalcinosis. Selection in follow-up of patients did not find place, other than by the willingness of children and their parents to participate in the study. This is reflected by the finding that children participating did not differ in gestational age or in birth weight from those not participating in the study. Children classified as having dubious neonatal nephrocalcinosis were not included in the follow-up study to get a clear picture of the influence of nephrocalcinosis on renal function.

### Blood Pressure

Blood pressure in our study did not differ in ex-preterm infants with and without nephrocalcinosis but was significantly higher than expected for healthy children, although only a minority (3 [7%] of 42 with and 2 [6%] of 31 without nephrocalcinosis) had systolic blood pressure more than the 95th percentile (Table 1). This is in concert with the current concept that high blood pressure is increasingly seen in follow-up of preterms into (young) adulthood and is a risk factor for hypertension later in life. Although there is evidence that low birth weight is an important factor in the development of hypertension and metabolic syndrome in adults, prematurity did but birth weight SDS did not predict hypertension in a large cohort of 19-year-old ex-preterm infants. Similarly, in our study there was no correlation between high blood pressure and birth weight SDS. Concluding, in our study, prematurity per se is associated with high blood pressure.

### Renal Ultrasound

Long-term follow-up data on nephrocalcinosis in ex-preterm infants are scarce. Studies in a limited number

---

**TABLE 2** Long-term Tubular Function and Urinary Solute Excretion of Premature Infants With (NC+) and Without (NC-) Neonatal Nephrocalcinosis

<table>
<thead>
<tr>
<th>Proximal tubular function</th>
<th>n</th>
<th>Median (Range)</th>
<th>n</th>
<th>Median (Range)</th>
<th>p&lt;sup&gt;a&lt;/sup&gt;</th>
<th>NC+ Abnormal, n (%)</th>
<th>NC- Abnormal, n (%)</th>
<th>p&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Reference Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRP, %</td>
<td>39</td>
<td>89 (69–100)</td>
<td>32</td>
<td>92 (77–100)</td>
<td>.02&lt;sup&gt;d&lt;/sup&gt;</td>
<td>7 (18)</td>
<td>0 (0)</td>
<td>.82&lt;sup&gt;c&lt;/sup&gt;</td>
<td>≥85</td>
</tr>
<tr>
<td>Urine glucose, g/L</td>
<td>39</td>
<td>—</td>
<td>32</td>
<td>—</td>
<td>.001&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>.30&lt;sup&gt;c&lt;/sup&gt;</td>
<td>&lt;0.3</td>
</tr>
</tbody>
</table>

Distal tubular function

| Plasma bicarbonate, mmol/L | 42 | 23 (20–28)    | 32 | 24 (21–32)    | .006<sup>d</sup> | 5 (12)            | 3 (9)             | .01<sup>d</sup> | 22–29         |
| Urine osmolality, mOsm/kg  | 41 | 880 (566–1123)| 32 | 931 (363–1223)| .15               | 9 (22)            | 6 (19)            | <0.001         | >700          |

Urinary excretion

| Urine calcium/creatinine, mmol/mmol | 41 | 0.24 (0.01–1.21) | 32 | 0.20 (0.01–0.76) | .15 | 9 (22) | <0.001 | 2 (6) | <0.06 |
| Urine calcium/citrate, mmol/mmol  | 39 | 0.94 (0.05–3.30) | 32 | 0.65 (0.04–1.6) | .007<sup>d</sup> | — | 0 (0) | <0.001 | <0.03 |
| Urine citrate/creatinine, mmol/mmol | 39 | 0.33 (0.07–0.65) | 32 | 0.38 (0.05–0.94) | .13 | 14 (36) | <0.001 | 12 (38) | <0.001 |

**TABLE 3** Kidney Length Related to Patient Height at Long-term Follow-up of Ex-preterm Infants With (NC+) and Without (NC-) Neonatal Nephrocalcinosis

<table>
<thead>
<tr>
<th>Percentile</th>
<th>NC+, n (%)</th>
<th>NC-, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEFT KIDNEY</td>
<td>RIGHT KIDNEY</td>
<td>LEFT KIDNEY</td>
</tr>
<tr>
<td>&lt;5th</td>
<td>1 (2.5)</td>
<td>0</td>
</tr>
<tr>
<td>5th–50th</td>
<td>30 (73)</td>
<td>27 (66)</td>
</tr>
<tr>
<td>51st–95th</td>
<td>10 (24.5)</td>
<td>14 (34)</td>
</tr>
<tr>
<td>&gt;95th</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>41 (100)</td>
<td>41 (100)</td>
</tr>
</tbody>
</table>

Left kidney length of NC+ compared with NC-: P = .04; NC+ left kidney: P = .004; NC+ right kidney: P = .04; NC- left kidney: P = .02; NC- right kidney: P = .05.

<sup>a</sup> Compared with NC- with healthy children.

<sup>b</sup> Comparison of NC+ with healthy children.

<sup>d</sup> P < .05.

Conversion rates: 1 mmol/mmol calcium/creatinine = 0.35 mg/mg; 1 mmol/mmol calcium/citrate = 0.21 mg/mg; 1 mmol/mmol citrate/creatinine = 1.7 mg/mg.
of patients show persistence of nephrocalcinosis in 4 of 10, 5 of 9, 6 of 26, 5 of 11, 4 of 16, 2 of 26, and 3 of 12 after 1 to 2 years, 21 ± 15 months, 2 years, 4 to 5 years, 3 to 6 years, 6 years, and 5.8 to 7.7 years, respectively. We observed only 10% nephrocalcinosis in ex-preterm infants after a longer follow-up period of 7.4 (±1.0) years. This fits in well with our previous findings of 34% and 15% persistence of nephrocalcinosis after 15 and 30 months, and the concept of spontaneous resolution of nephrocalcinosis in time.

Furthermore, we established that kidneys of patients with as well as without neonatal nephrocalcinosis were significantly smaller than expected for healthy children of the same height, with the exception of left kidneys of children without neonatal nephrocalcinosis, which showed only a trend in being smaller (P = .05). This is in concert with the study of Keijzer-Veen, who found decreased kidney growth in 51 preterms (gestational age: <32 weeks) compared with 30 term born controls at the age of 20 years. It is hypothesized that preterm neonates are specifically at risk of renal growth impairment because nephrogenesis peaks at 32 weeks and continues until 36 weeks.

Glomerular Function

Undoubtedly, the most important question is, does nephrocalcinosis in very prematurely born children affect growth and function of kidneys in the long-term? To date, only small numbers of patients have been studied. Furthermore, most studies probably concern selected populations of prematurely born children. The first study to address this question was performed by Ezedeen et al. Four of 9 children with nephrocalcinosis as preterm neonate had low GFR at a mean age of 21 (SD: 15) months (20, 25, 52, and 65 mL/min per 1.73 m², respectively). At 2 years, Downing et al. also found significantly lower mean GFR in 10 children with nephrocalcinosis as preterm neonates (84 ± 8 mL/min per 1.73 m²) compared with 2 groups without nephrocalcinosis but with (n = 10; 109 ± 5 mL/min per 1.73 m²) or without furosemide medication (n = 7; 103 ± 7 mL/min per 1.73 m²) in the neonatal period. Jones et al. studied 11 children at the age of 4 to 5 years with nephrocalcinosis as preterm neonates. She found a median GFR of 61 mL/min per 1.73 m² (range: 46–79 mL/min per 1.73 m²), which is low compared with reference values for healthy children. This is in contrast with Saarela et al., who found no significant difference in GFR at a mean age of 4.7 years between 20 preterms with and without nephrocalcinosis in the neonatal period. Hoppe et al. and Porter et al. also observed normal GFR after 3 to 6 and 5.8 to 7.7 years, respectively, in 12 and 14 prematurely born children with nephrocalcinosis.

We previously reported mean estimated GFR of 110 (SD: 34) and 132 (SD: 34) mL/min per 1.73 m² after 1 and 2 years in the same population of preterm neonates with nephrocalcinosis studied here. At that time, none of the children at 1 year and only 2% of the children at 2 years of age had low values for GFR compared with healthy subjects. Here, we extended the follow-up period to mean 7.4 ± 1 years and, in addition, studied preterm neonates without neonatal nephrocalcinosis as a control group. Significantly more children in the nephrocalcinosis group (6 of 40 [15%]) had low GFR compared with healthy children; this is in contrast to children without neonatal nephrocalcinosis (2 of 32 [6%]; Fig 1). However, there was no significant difference in GFR or microalbuminuria between both groups. Interestingly, the degree of nephrocalcinosis at term (focal or extensive) did not influence glomerular function at follow-up. Furthermore, no correlation between persistence of nephrocalcinosis and low GFR was found, but the number ex-preterm infants with persisting nephrocalcinosis (n = 4) is small. We conclude that long-term follow-up of preterm neonates with nephrocalcinosis demonstrates an unfavorable effect on renal function in some children.

Proximal Tubular Function

If neonatal nephrocalcinosis in preterms has long-term sequelae for glomerular function, how does nephrocalcinosis affect tubular function? Jones et al. studied 11 ex-preterm infants at the age of 4 to 5 years with neonatal nephrocalcinosis. She found a low median tubular phosphate reabsorption/100 mL glomerular filtration rate compared with reference values. Likewise, Saarela et al. also observed a significantly higher urine β2-microglobulin/creatinine ratio in 20 children with compared with 20 children without nephrocalcinosis as preterm neonates at a mean age of 4.7 years. However, in his study TRP did not differ significantly in children with and without nephrocalcinosis. In contrast, Downing et al. found significantly lower TRP (84 ± 2% vs 93 ± 1%; P < .05) in preterm children at the age of 1 to 2 years with nephrocalcinosis compared with preterms without nephrocalcinosis. In our study, TRP also was significantly lower in children with compared with without nephrocalcinosis (Table 2). However, plasma phosphate was within reference limits in all children. Therefore, the implication of low TRP in these patients is debatable. We did not observe glucosuria in any of the children.

These findings are in concert with our previous study where we found normal tubular phosphate reabsorption/100 mL glomerular filtration rate and urine α1-microglobulin in ex-preterm infants with neonatal nephrocalcinosis after 1 and 2 years follow-up. Considering these data, our study does not support firm evidence for proximal tubular dysfunction caused by neonatal nephrocalcinosis in ex-preterm infants.
Distal Tubular Function

Downing et al.\textsuperscript{36} observed lower ability to excrete hydrogen ions in the distal tubule in preterms with ($n = 10$) compared with without ($n = 14$) nephrocalcinosis. In another study, distal tubular acidification was measured by using an oral acetazolamide test, in which the response was abnormal in 1 of 20 children with nephrocalcinosis as opposed to none without neonatal nephrocalcinosis.\textsuperscript{25} On the other hand, Hoppe et al.\textsuperscript{7} noted no acidosis after a follow-up period of 3 to 6 years in 12 preterms with neonatal nephrocalcinosis. In our study, we found median plasma bicarbonate was significantly lower in children with versus without neonatal nephrocalcinosis (Table 2). The urine anion gap of the children with low plasma bicarbonate was inappropriately high, indicating distal rather than proximal tubular dysfunction.

Downing et al.\textsuperscript{36} found no difference in early-morning urine osmolality between ex-preterm infants with and without nephrocalcinosis after 1 to 2 years. In another study, 12 preterm infants with neonatal nephrocalcinosis had early-morning urine osmolality within the reference range after a follow-up period of 3 to 6 years.\textsuperscript{7} Likewise, Porter et al.\textsuperscript{22} noted early-morning urine osmolality that was within the reference range in 14 children with and 14 children without neonatal nephrocalcinosis at a mean age of 6.9 years. However, we previously found impaired desmopressin test in 4 of 30 ex-preterm infants with neonatal nephrocalcinosis at 1 year and 2 of 25 at 2 years.\textsuperscript{21} In our current study, early-morning urine osmolality did not differ between children with and without neonatal nephrocalcinosis, but likewise was significantly lower in both groups compared with healthy children. However, it is difficult to rule out that some of the children were not fasting when collecting the early-morning urine sample. In concert with our study, Jones et al.\textsuperscript{24} found no significant difference in renal concentrating capacity after desmopressin in children with and without nephrocalcinosis. Equally, she found the mean concentrating capacity was below that of reference values.

Interpreting all these data, nephrocalcinosis in preterm neonates can have long-term sequelae mainly for distal tubular function. However some tubular defects cannot solely be attributed to nephrocalcinosis but are also seen in ex-preterm infants without neonatal nephrocalcinosis (low early-morning urine osmolality and plasma bicarbonate).

Hypercalciuria

We found significantly more (9 of 41 [22%]) children with hypercalciuria in the group with neonatal nephrocalcinosis than expected in healthy children; this is in contrast to ex-preterm infants without nephrocalcinosis (2 of 32 [6%]). In addition, the urinary calcium/citrate ratio was significantly higher at long-term follow-up in children with compared with without neonatal nephrocalcinosis (Table 2). Hypercalciuria is indeed frequently seen in ex-preterm infants with neonatal nephrocalcinosis. At 1 to 2 years of age, the urine calcium/creatinine ratio was significantly higher in 10 preterms with (0.94 ± 0.23 mmol/mmol) compared with 10 preterms without (0.45 ± 0.17 mmol/mmol) nephrocalcinosis.\textsuperscript{26} Similarly, the urine calcium/creatinine ratio was significantly higher in 20 children with (0.68 ± 0.45 mmol/mmol) versus 20 children without (0.34 ± 0.23 mmol/mmol) neonatal nephrocalcinosis at a mean age of 4.7 years.\textsuperscript{25} On the other hand, Jones et al.\textsuperscript{24} found no difference in calcium excretion in 7- to 8-year-old ex-preterm infants compared with healthy controls.\textsuperscript{37} Porter et al.\textsuperscript{22} also noted evidence of hypercalciuria in 2 of 14 children with as well as in 4 of 14 children without neonatal nephrocalcinosis at long-term follow-up, suggesting prematurity might be a risk factor.

Furthermore, in our study, the urine citrate/creatinine ratios in both groups were significantly lower than expected in healthy children (reference value: >0.3 mmol/mmol, 0.51 g/g). Then again, there is no clear consensus in the literature on the cutoff level for healthy children (Table 2).\textsuperscript{38}

CONCLUSIONS

Nephrocalcinosis in preterm neonates can have long-term sequelae for glomerular and tubular function (low plasma bicarbonate, high urine calcium/citrate ratios). Furthermore, prematurity per se is associated with high blood pressure and relatively small kidneys, as well as (distal) tubular dysfunction (low plasma bicarbonate and early-morning urine osmolality). Long-term follow-up of blood pressure and renal function of prematurity born children, especially with neonatal nephrocalcinosis, seems warranted. Future research pertaining to prevention of nephrocalcinosis in preterm neonates is needed.

ACKNOWLEDGMENT

We gratefully acknowledge financial support by the Dutch Kidney Foundation.

REFERENCES


Steroid Use Among Adolescents: Longitudinal Findings From Project EAT

Patricia van den Berg, PhD, Dianne Neumark-Sztainer, PhD, RD, MPH, Guy Cafri, MA; Melanie Wall, PhD

Divisions of Epidemiology and Community Health and Biostatistics, School of Public Health, University of Minnesota, Minneapolis, Minnesota; Department of Psychology, University of South Florida, Tampa, Florida

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

ABSTRACT

OBJECTIVE. We examined the prevalence, persistence, secular and longitudinal trends, and predictors of steroid use in a diverse sample of adolescents.

PARTICIPANTS AND METHODS. Data are from Project EAT-II (Eating Among Teens), a 5-year longitudinal study of eating, activity, weight, and related variables in 2516 middle and high school students. Data were collected in 1999 (time 1) and 2004 (time 2).

RESULTS. Approximately 1.5% of adolescents reported steroid use at time 2. Use differed by ethnicity but not socioeconomic status. Steroid use was not stable across time, although the risk of use at time 2 was higher for girls and (marginally) for boys who used steroids at time 1. No secular trends were noted in middle adolescents’ steroid use between 1999 and 2004. Developmentally, steroid use decreased as adolescents grew older. Predictors of use for male adolescents included wanting to weigh more and reporting higher use of healthy weight-control behaviors. Female time 2 steroid users had higher BMIs and were less satisfied with their weight, had poorer nutrition knowledge and concern for health, and were marginally more likely to have participated in weight-related sports at time 1.

CONCLUSIONS. The prevalence of steroid use in adolescents was low but of concern. Although use was not persistent over 5 years, time 1 use was a risk factor for time 2 use in female adolescents. There was no change in the prevalence of steroid use by middle adolescents between 1999 and 2004 despite a great deal of public interest in steroids during this time period. Steroid use decreased as adolescents grew older. Weight-related variables predicted adolescents’ steroid use 5 years later, and health and nutrition knowledge and concern and (marginally) participation in weight-related sports further predicted use in female adolescents. These findings suggest that early preventive efforts may be most useful.
**ANABOLIC-ANDROGENIC STEROIDS ARE synthetic derivatives of testosterone that act to increase protein synthesis and the development of male secondary sex characteristics.** They are typically taken to increase muscle mass and strength, for either improved sports performance or to enhance appearance. In studies of adults, anabolic steroids have been found to have significant adverse effects on the musculoskeletal, cardiovascular, endocrine/reproductive, and hepatic systems, as well as variability in mood and other possible psychological effects. Given these serious consequences, it is important to understand the prevalence, persistence, secular changes, longitudinal trends, and predictors of steroid use in adolescents.

**PREVALENCE**
The prevalence of steroid use has been estimated in several large survey studies of adolescents. Overall, a nontrivial percentage of adolescents admit having used steroids, with boys generally having higher rates. The 2004 Monitoring the Future Study reported that annual prevalences were 1.3% and 3.3% for 8th- and 12th-grade boys, respectively, whereas the prevalences in girls were 1.0% and 1.7% in 8th- and 12th-graders. These prevalences were based on samples of ~17 000 8th-graders and 14 600 12th-graders. Other surveys have shown comparable prevalences. For instance, in a previous cross-sectional analysis of the time 1 data we used in this study (N = 4476), we found annual prevalences of 4.4% among older adolescent boys and 1.4% among older adolescent girls, and slightly higher prevalences of 7.6% (boys) and 5.7% (girls) among younger adolescents. The 2005 Youth Risk Behavior Surveillance (YRBS) study reported lifetime prevalences of 4.8% in boys and 3.2% in girls in their sample of 13 953 adolescents in grades 9 through 12.

**SECULAR TRENDS**
The prevalence of steroid use among adolescents rose throughout the 1990s, causing concern among health professionals. However, research has not consistently shown that this increase continued into the next century. For instance, the Monitoring the Future study found sharp increases in steroid use in 1999–2000, especially among boys. However, prevalences for the most part leveled off thereafter. In the YRBS, there was an upward trend in lifetime prevalence of steroid use between 1991 and 2003, but between 2003 and 2005 prevalences decreased somewhat. Given recent publicity surrounding the use of anabolic-androgenic steroids and other performance-enhancing substances by professional and elite athletes, it is important to determine whether use is still increasing or has leveled off or even decreased over the last 5 to 10 years.

**LONGITUDINAL TRENDS**
The research to date on developmental changes in steroid use is contradictory and primarily uses data from cross-sectional studies, which are not ideal for investigating longitudinal trends. In Monitoring the Future, a cross-sectional study, prevalence of steroid use seemed to rise across increasing grade level. However, other samples have shown decreasing rates with increasing age or grade. For instance, a survey of Connecticut youth found decreases in use across 7th to 11th grades in male adolescents. Regarding longitudinal studies, most research has involved constructs related to steroid use, but not steroid use itself. For instance, a recent longitudinal study of preadolescents used a composite measure of different strategies to increase muscularity, including exercising, eating, and the use of food supplements, but not the use of steroids. Findings indicated that strategies to gain muscles decreased over 16 months of follow-up in 8- to 11-year-olds. In a similar study of adolescent boys 11 to 16 years old, researchers found a correlation of only 0.37 over 8 months on scores on a composite measure of strategies to increase muscles. The only longitudinal study of steroid use itself that we could locate was conducted recently by Dodge and Jaccard, who used data on 15 000 adolescents in the National Longitudinal Study of Adolescent Health. These researchers examined longitudinal predictors of steroid use but did not investigate longitudinal trends in the prevalence of steroid use. Although the studies conducted to date are suggestive, research that specifically examines longitudinal trends in steroid use is necessary to understand developmental changes in steroid use in adolescents.

**PREDICTORS OF STEROID USE**
Previous studies have identified several possible predictors of anabolic steroid use. One cross-sectional study of middle school boys found that self-reported media and parent and peer pressures regarding weight or muscles, as well as depression, negative body image, and tendency to compare one’s appearance to that of others all distinguished boys who used steroids from those who did not. Other researchers have found that perceived pressure to increase muscularity from the media, parents, and peers, as well as increased negative or decreased positive affect, predicted the use of muscle-building strategies 16 months later in both boys and girls; however, the authors did not study steroid use specifically. In previous cross-sectional studies of body image and steroid use, a complex relationship has been found, suggesting perhaps that wanting to be larger may be associated with onset of use, but that users may have increased satisfaction. However, longitudinal studies of body image and steroid use have not been conducted to date. Participation in various types of sports, especially power sports such as football, wrestling, and track and
field, have been found in cross-sectional studies to be associated with increased steroid use, especially in boys.\textsuperscript{2,14,15} In addition, the study by Dodge and Jaccard described above examined the longitudinal association between high school sports participation and later steroid use. They found a nonsignificant association between sports participation and steroid use, but a significant interaction between gender and sports participation, such that boys were at higher risk for steroid use overall compared with girls, and even more so if they had participated in sports in high school. The cross-sectional findings regarding higher or lower BMI as a possible risk factor have been inconsistent across age group and study, with some studies finding that lower BMI was associated with steroid use, and others finding no relationship, or a relationship with higher BMI.\textsuperscript{2,5} Steroid use in boys has also been associated in several cross-sectional studies with the use of other illicit substances and other high risk behaviors, such as having unprotected sex, driving while under the influence of alcohol, and possessing a gun.\textsuperscript{14,16} In sum, cross-sectional studies have generated a number of suggestive associations, and 1 longitudinal study has suggested that sports participation predicts steroid use longitudinally, at least in boys. However, we are unaware of any longitudinal studies that have tested associations with other variables over time.

Our group previously examined correlates of steroid use in a cross-sectional analysis of the time 1 data from Project EAT (Eating Among Teens\textsuperscript{5}), which is the project on which the current longitudinal study is based. We found that for boys, factors associated with higher steroid use included dissatisfaction with one’s shoulders, parental concern and family teasing about weight, self-report of bingeing and unhealthy weight-control behaviors and self-report of having received an eating disorder diagnosis, low self-esteem, depressed mood and a history of suicide attempts, poorer health and nutrition knowledge and attitudes, participation in weight-related sports, and use of alcohol, cigarettes, marijuana, and other drugs. The significant correlates of steroid use were similar for girls, including weight and shape concerns, low self-esteem and a history of suicide attempts, poorer nutrition and health knowledge and attitudes, involvement in weight-related sports, parental concern about weight, unhealthy weight-control behaviors, binging and eating disorder diagnosis, and use of marijuana and other drugs. However, because this was a cross-sectional study, the temporal order of these associations could not be established.

In summary, nearly all of the literature on steroid use in adolescence is based on cross-sectional studies and, therefore, cannot inform our understanding of developmental changes in steroid use and cannot distinguish between predictors and correlates of steroid use. Our study is a 5-year longitudinal study intended to address the gaps in the literature on steroid use in adolescence and to extend the findings of our previous work with this population. The aims of this study were to examine the prevalence of steroid use in our sample in 2004–2005 (time 2), the persistence of steroid use between time 1 and time 2, the secular and longitudinal changes in steroid use across the 5 years of follow-up, and the personal, socioenvironmental, and behavioral variables that predicted steroid use 5 years later.

METHODS

Study Sample and Design

Project EAT is a 5-year longitudinal study examining eating behaviors, weight concerns, and related variables in a large, ethnically diverse population of adolescents.\textsuperscript{17,18} The first wave of data collection, time 1, took place in the 1998–1999 academic year, and the second wave, time 2, took place in the 2003–2004 academic year. Participants were recruited from 31 urban and suburban public middle and high schools in the Minneapolis/St Paul, Minnesota, metro area. The study sample at time 1 consisted of 4746 adolescents in 7th through 12th grades, with approximately equal numbers of boys and girls. Participants completed in-class surveys, and trained research staff measured their height and weight in a private area of the school. At time 2, surveys were mailed to the address given by the participant at time 1. In those cases where mail was returned because of an incorrect address, Internet tracking services were used to identify correct addresses. Initial nonresponders were sent 2 reminder postcards and 3 additional survey packets to encourage participation. Of the original sample at time 1, 1074 (22.6%) were unable to be followed, primarily because of missing or obsolete contact information. Of the remaining 3672 participants to whom surveys were mailed, 2516 responded, comprising 53.0% of the original cohort and 68.4% of participants with valid contact information at time 2. Of the 186 steroid users at time 1, 38 boys (32% of the time 1 male users) and 33 girls (49% of the time 1 female users) were included in our study.

The final study population consisted of 1130 boys (45%) and 1386 girls (55%) who completed surveys at both time 1 and time 2. The one third of the participants (32%) who were originally in middle school comprised the younger cohort; at time 1 their mean age was 12.8 years (SD: 0.8) and at time 2 their mean age was 17.2 years (SD: 0.6). The two thirds of the participants (68%) who were originally recruited from high schools constituted the older cohort; at time 1 their mean age was 15.8 years (SD: 0.8) and at time 2 their mean age was 20.4 years (SD: 0.8). Of this older cohort, 66% reported attending school full- or part-time during the previous year, 48% reported that they lived in their parents’ home during the previous year, and 54% reported that they worked <30 hours per week during the previous
BMI was computed from measured height and reported current weight and multiplying this ratio by weight do you think you would look best?" by self-dividing self-reported “ideal weight” (ie, “At what you feel about yourself?” Three items from the Body Shape Satisfaction scale23 were used to assess satisfaction with appearance. Participants rated on a 5-point scale their level of satisfaction with their body build, shoulders, and weight. Cronbach’s α for the entire 10-item scale in this study was .92. Weight discrepancy was calculated by subtracting self-report from measured weight. Cronbach’s α at time 1 was .76. Weight importance was assessed with an item adapted from the Body Shape Satisfaction scale23: “During the past six months, how important has your weight or shape been in how you feel about yourself?” Three items from the Body Shape Satisfaction scale23 were used to assess satisfaction with appearance. Participants rated on a 5-point scale their level of satisfaction with their body build, shoulders, and weight. Cronbach’s α for the entire 10-item scale in this study was .92. Weight discrepancy was calculated by dividing self-reported “ideal weight” (ie, “At what weight do you think you would look best?”) by self-reported current weight and multiplying this ratio by height in meters squared.

Psychological Measures
A 6-item shortened version of the Rosenberg Self-esteem scale25 assessed general self-esteem. Cronbach’s α for the scale was .79. Depressive symptoms were assessed using a 6-item scale developed by Kandel and Davies.26 Chronbach’s α in the current sample was .82 at time 1. Suicidal ideation and previous attempts were each assessed with 1 question: “Have you ever thought about killing yourself?” and “Have you ever tried to kill yourself?” These questions have been used in previous population-based surveys of adolescents.27

Health/Nutrition Knowledge and Attitudes
To assess knowledge of healthy eating, participants were instructed to indicate which food was healthier for each of 7 pairs of food items (for example, “frozen yogurt” and “ice cream”). The participant’s score was the number of correct responses. Cronbach’s α for this scale was .63. To assess self-efficacy for making healthy food choices, participants indicated their level of certainty that they could eat healthy foods when feeling certain emotions (for example, “stressed out”) or when in certain situations (for example, “hungry after school”). The scale consisted of 9 items, and Cronbach’s α was .85. Concern about health was assessed by responses to 5 questions regarding the importance of health and healthy eating. For this measure, Cronbach’s α was .70 at time 1.

Socioenvironmental Factors
Sports Involvement
Weight related sports participation was assessed with the question “Are you in a sport or activity where it’s important to stay a certain weight (wrestling, gymnastics, ballet, etc)?”

Weight-Related Norms/Teasing
Parental concern with weight consisted of 4 items asking about the participant’s mother’s and father’s tendency to diet and to encourage the participant to diet to control weight. Cronbach’s α at baseline was .76. Peer dieting behavior was assessed by the question, “Many of my friends diet to lose weight or keep from gaining weight.” Weight teasing by family and by peers were each assessed with 1 question adapted from the Perception of Teasing scale, which has shown good reliability and validity in previous studies.28

Behavioral Factors
Physical Activity Level
A modified version of the Leisure Time Exercise Questionnaire29 was used to assess hours of weekly activity. Participants reported the number of hours per week spent engaging in strenuous, moderate, and mild exercise. The test-retest correlation for this measure was .69 in previous research.30

Unhealthy Eating/Weight Control
To assess weight-control behaviors, participants were instructed to endorse the methods they had used during the past year to lose or maintain weight. Four items were healthy weight-control behaviors (for example, “ate more fruits and vegetables”), and 5 were unhealthy
behaviors (for example, “skipped meals”). To assess current weight gain/loss attempts, participants indicated whether they were currently trying to lose, maintain, or gain weight, or not doing anything about their weight. Self-report of eating disorder diagnosis was assessed with the question “Has a doctor ever told you that you have an eating disorder such as anorexia nervosa, bulimia nervosa, or binge eating disorder?” Binge eating was assessed with a combination of 2 items regarding binging and loss of control, from the Questionnaire on Eating and Weight Patterns-Revised. Participants who answered “yes” to both questions were classified as having binged.

Substance Use

Participants were asked about past year use of cigarettes, alcohol, marijuana, and other drugs (“drugs other than marijuana [acid, crack, cocaine, etc]”). Response options ranged from “never” to “daily,” and each item was dichotomized for analysis (“never” or “any”).

Sociodemographic Characteristics

Gender, age (in years), ethnicity/race, and socioeconomic status (SES) were based on self-report at time 1. SES was calculated using an algorithm that weighted parental education level most heavily but also took into account family eligibility for public assistance, eligibility for free or reduced-cost school meals, and employment status of the mother and father (see ref 18 for more information).

Statistical Analysis

Attrition at time 2 differed across sociodemographic characteristics. Thus, in all analyses participants’ responses were weighted to adjust for this differential response rate. The response propensity method was used to generate the weights, such that an individual’s weight was the inverse of the estimated probability that the individual responded at time 2. Response propensities (ie, the probability of responding to the time 2 survey) were estimated using a logistic regression, with a large number of predictor variables from the time 1 survey predicting response at time 2 (yes/no). The selected response propensity model included main effects for time 1 gender, native-born status, ethnicity/race, SES, overweight status, parental marital status, individual’s concern about health, and most common grade received in school. In addition, weights were calibrated so that the weighted total sample sizes used in analyses for each gender cohort accurately reflected the actual observed sample sizes in those groups. The weighting method resulted in estimates representative of the demographic make-up of the original time 1 sample. The weighted ethnic/racial and SES proportions are as follows: 48.3% white, 18.9% black, 5.8% Hispanic, 19.6% Asian, 3.6% Native American, and 3.8% mixed or other race, and SES was low (17.8%), middle-low (18.9%), middle (26.7%), middle-high (23.3%), and high (13.3%).

All analyses were stratified by gender. Initial analyses included examination of the frequency of steroid use in male and female adolescents and demographic characteristics of adolescents who did and did not use steroids. Persistence of steroid use was investigated by examining the percentages of participants who used steroids at time 1 and time 2, as well as the incidence of new users of steroids and time 1 users who did not report use at time 2. Secular and longitudinal trends in prevalence of steroid use were examined using mixed model regressions, with main effects for both year (1999 or 2004) and cohort (younger or older), a year by cohort interaction, and a random effect for individuals to account for longitudinal correlation. These mixed models were stratified by cohort and gender and adjusted for race/ethnicity and SES. Adjustment for race/ethnicity and SES was conducted to assure that the comparisons between 1999 and 2004 prevalences were not confounded by differences between the older and younger cohorts on race/ethnicity and SES.

Logistic regression was used to estimate the association between time 1 variables and the outcome of steroid use. Unadjusted means and prevalences for steroid users and nonusers are reported for continuous and dichotomous independent variables, respectively. Odds ratios (ORs) and the corresponding P values are adjusted for time 1 steroid use and sampling weights (although not for demographics because of the small number of steroid users). Standardized odds ratios are presented to allow comparison across differently scaled variables. P values were set at .05 without adjustment for multiple tests because of the exploratory nature of these analyses and the low base rate of steroid use. SAS 9.1 was used for all analyses.

RESULTS

Descriptive Characteristics

Overall, 1.4% of female and 1.7% of male adolescents reported having used anabolic steroids in the last year; the difference in prevalence between genders was not statistically significant (P = .57) (Table 1). Among the male steroid users, 54% reported having used steroids a few times in the past year, 18% reported monthly use, and 10% reported daily use. The percentages for female users were 83%, 5%, 12%, and 0%, respectively. Regarding differences by age cohort, the younger group showed a higher prevalence of steroid use, which in male adolescents reached significance. Boys in the younger cohort, which had been in middle school at time 1, were nearly 3 times more likely to report steroid use than male adolescents in the older cohort.

The prevalence of steroid use differed across race/
TABLE 1  Time 2 Anabolic Steroid Use to Gain Muscle in the Past 12 Months According to Gender, Cohort, Race, and SES

<table>
<thead>
<tr>
<th></th>
<th>Male Adolescents (n = 1130)</th>
<th>Female Adolescents (n = 1386)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1.7</td>
<td>1.4</td>
</tr>
<tr>
<td>Female</td>
<td>1.9</td>
<td>1.8</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Older</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Younger</td>
<td>3.1</td>
<td>1.9</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1.3</td>
<td>1.0</td>
</tr>
<tr>
<td>Black</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Other Asian</td>
<td>1.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Native American</td>
<td>9.8</td>
<td>0.0</td>
</tr>
<tr>
<td>Mixed</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Hmong</td>
<td>4.2</td>
<td>2.8</td>
</tr>
<tr>
<td>SES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>2.2</td>
<td>0.8</td>
</tr>
<tr>
<td>Low-middle</td>
<td>1.4</td>
<td>2.3</td>
</tr>
<tr>
<td>Middle</td>
<td>2.5</td>
<td>1.1</td>
</tr>
<tr>
<td>High-middle</td>
<td>0.6</td>
<td>2.3</td>
</tr>
<tr>
<td>High</td>
<td>1.6</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Frequencies were weighted for nonresponse at time 2.  
*P*-values are from chi2 tests of independence.

Temporal Stability of Self-reported Steroid Use

The temporal stability of steroid use was low overall. Of the 41 boys who reported steroid use at time 1, only 2 (4%) also reported use at time 2 (Table 2). Of 45 girls who indicated they had used steroids at time 1, only 5 (11%) again indicated that they used steroids at time 2. The incidence of new users was 1.3% for boys and 1.2% for girls. The odds of using steroids at time 2 were not significantly higher for boys who used steroids at time 1 compared with nonusers at time 1 (P = .15; OR: 3.15; 95% confidence limit: 0.60, 16.51). However, for girls the odds of time 2 use were ~10 times higher among those who had used at time 1, compared with those who were nonusers at time 1 (OR: 10.37; 95% confidence limit: 3.57, 30.17).

Secular and Longitudinal Trends in Steroid Use

Figure 1 depicts the secular trends in steroid use for middle adolescent boys and middle adolescent girls between 1999 and 2004, as well as the developmental changes in steroid use between early and middle adolescence, and between middle and late adolescence. Secular changes can be identified by comparing the prevalence in the younger cohort (solid lines) to that in the older cohort (broken lines) at the middle of the graph, which is middle adolescence. In neither male (difference in prevalence: 1.7%, P = .14) nor female adolescents (difference in prevalence: 0.1%, P = .91) was there a significant secular trend in steroid use between 1999 and 2004 among middle adolescents.

However, both the younger and older cohorts of male adolescents showed significant longitudinal decreases in steroid use across the 5 years of follow-up. Among female adolescents, only the younger cohort had a significant longitudinal decrease in steroid use between time 1 and time 2. The older female cohort had a very low prevalence at time 1, creating a floor effect with very little decrease possible over follow-up.

Predictors of Steroid Use

For male adolescents, 2 variables were significant predictors of time 2 steroid use. Having an ideal body size that is larger than one’s current body size and self-report of healthy weight-control behaviors at time 1 predicted steroid use 5 years later (Table 3).

For female adolescents, lower satisfaction with weight and higher BMI at time 1 were significant predictors of time 2 steroid use (Table 4). Time 2 steroid users also showed a higher mean level of satisfaction with their shoulders at time 1. In addition, lower levels of knowledge of healthy eating and concern with health at time 1 characterized time 2 users. Those who used steroids at time 2 were over twice as likely as nonusers to have participated in weight-related sports at time 1, although this trend was only marginally significant.

DISCUSSION

The aims of this study were to examine the prevalence of steroid use, the persistence of steroid use across time,
secular and longitudinal changes in steroid use over 5 years, and the personal, environmental, and behavioral predictors of steroid use 5 years later. This study comprises one of the first longitudinal investigations of steroid use in adolescents.

In this second wave of Project EAT, ~1.5% of male and female adolescents reported having used steroids for muscle gain, with no significant difference in prevalence between genders. With regard to the persistence of steroid use across the 5 years of follow-up, the vast majority of participants who were using at time 1 were not still using at time 2, although the odds of time 2 use were 10 times higher for girls who had used at time 1 compared with nonusers at time 1. Likewise, although not statistically significant, the odds of using at time 2 were 3 times higher for boys who had reported use at time 1. There were no significant secular changes found for steroid use between 1999 and 2004. However, in our examination of longitudinal trends we found that the prevalence of steroid use decreased significantly as adolescents grew older for all but the older female cohort. Regarding the longitudinal predictors of steroid use, few variables held up over 5 years and after adjusting for baseline steroid use. For boys, wanting to have a larger body size predicted steroid use 5 years later, as did the use of healthy weight-control behaviors. For girls, lower satisfaction with body weight and higher BMI predicted later steroid use, as did lower levels of knowledge of healthy eating and concern with health. There was also a marginally significant relationship between time 2 steroid use and time 1 participation in weight-related sports. Higher satisfaction with their shoulders also predicted time 2 use among girls.

The prevalences of steroid use in this study are similar to, or lower than, those reported in other studies. For instance, as previously discussed, the 2004 Monitoring the Future Study reported annual prevalences of 1.3% and 3.3% of middle and high school boys, and 1.0% and 1.7% of middle and high school girls, in line with our findings. It is possible that the prevalences found in this study are lower than some others due to underreporting, because complete anonymity cannot be maintained in a longitudinal study. However, if this were biasing the prevalences, we would have expected to see lower time 1 prevalences as well, which we did not. The lack of a gender difference in this study, however, is decidedly different from other studies, most of which consistently found that more male than female adolescents use steroids. Some studies have reported increases in steroid use among female subjects during the 1990s.
thermore, the 2005 YRBS showed no difference between genders in lifetime prevalence of steroid use for individuals in 9th grade, although in 10th through 12th grade, boys did have higher prevalences.7 The prevalences in our study and these other studies may reflect a shrinking gender gap in adolescent steroid use, although there is some controversy surrounding the validity of recent self-reports of steroid use by girls.35

The fact that steroid use was not very persistent across the 5 years is consistent with other studies that found steroid use to be less persistent than the use of other substances.4 This seems to point more toward experimental use or to a pattern of alternating initiation and cessation of use rather than a pattern in which initiation leads to continued use, as seems to be the case with alcohol, for instance.4 Even so, the risk of future steroid use is greatly increased in female users at time 1.

Secularly, steroid use was stable in adolescents in high school between 1999 and 2004. This corroborates the findings of other studies, which have found preva-

### TABLE 3

<table>
<thead>
<tr>
<th>Time 1 Variable</th>
<th>Time 2 Nonusers (n = 1074)b</th>
<th>Time 2 Users (n = 19)b</th>
<th>OR</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight and shape concerns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight concerns, mean (SD)</td>
<td>5.9 (2.2)</td>
<td>6.1 (3.0)</td>
<td>1.04</td>
<td>.89</td>
</tr>
<tr>
<td>Weight importance, mean (SD)</td>
<td>1.9 (0.9)</td>
<td>2.2 (0.9)</td>
<td>1.37</td>
<td>.19</td>
</tr>
<tr>
<td>Satisfaction, body build, mean (SD)</td>
<td>3.6 (1.1)</td>
<td>3.4 (1.3)</td>
<td>0.79</td>
<td>.59</td>
</tr>
<tr>
<td>Satisfaction, shoulders, mean (SD)</td>
<td>3.9 (1.0)</td>
<td>3.7 (1.1)</td>
<td>0.86</td>
<td>.53</td>
</tr>
<tr>
<td>Satisfaction, weight, mean (SD)</td>
<td>3.5 (1.1)</td>
<td>3.2 (1.2)</td>
<td>0.80</td>
<td>.38</td>
</tr>
<tr>
<td>Ideal/estimated weight, mean (SD)</td>
<td>99.9 (15.6)</td>
<td>106.1 (18.2)</td>
<td>1.39</td>
<td>.04</td>
</tr>
<tr>
<td>BMI, mean (SD)</td>
<td>22.4 (4.5)</td>
<td>22.0 (5.2)</td>
<td>0.93</td>
<td>.78</td>
</tr>
<tr>
<td><strong>Psychological variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-esteem, mean (SD)</td>
<td>18.7 (3.5)</td>
<td>17.9 (2.9)</td>
<td>0.84</td>
<td>.50</td>
</tr>
<tr>
<td>Depressed mood, mean (SD)</td>
<td>9.7 (2.7)</td>
<td>8.7 (2.1)</td>
<td>0.55</td>
<td>.07</td>
</tr>
<tr>
<td>Suicide, thoughts, % (n)</td>
<td>18.7 (188)</td>
<td>7.8 (1)</td>
<td>0.38</td>
<td>.34</td>
</tr>
<tr>
<td>Suicide, attempt, % (n)</td>
<td>5.0 (50)</td>
<td>7.8 (1)</td>
<td>1.46</td>
<td>.72</td>
</tr>
<tr>
<td><strong>Health/nutrition attitudes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge, healthy eating, mean (SD)</td>
<td>3.7 (2.8)</td>
<td>2.3 (3.0)</td>
<td>0.71</td>
<td>.15</td>
</tr>
<tr>
<td>Efficacy, healthy food choices, mean (SD)</td>
<td>31.8 (9.7)</td>
<td>29.9 (5.2)</td>
<td>0.81</td>
<td>.44</td>
</tr>
<tr>
<td>Concern about health, mean (SD)</td>
<td>15.8 (2.4)</td>
<td>16.1 (1.9)</td>
<td>1.30</td>
<td>.35</td>
</tr>
<tr>
<td><strong>Socio-environmental factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sports involvement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight-related sports, % (n)</td>
<td>15.9 (163)</td>
<td>30.2 (5)</td>
<td>2.23</td>
<td>.18</td>
</tr>
<tr>
<td><strong>Weight-related norms/teasing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental concern with weight, mean (SD)</td>
<td>7.2 (3.0)</td>
<td>8.0 (3.9)</td>
<td>1.19</td>
<td>.48</td>
</tr>
<tr>
<td>Peer dieting, mean (SD)</td>
<td>2.3 (1.5)</td>
<td>2.3 (1.2)</td>
<td>0.97</td>
<td>.92</td>
</tr>
<tr>
<td>Teased about weight (family), % (n)</td>
<td>14.9 (150)</td>
<td>30.9 (5)</td>
<td>1.96</td>
<td>.27</td>
</tr>
<tr>
<td>Teased about weight (peers), % (n)</td>
<td>24.5 (248)</td>
<td>31.4 (5)</td>
<td>1.48</td>
<td>.49</td>
</tr>
<tr>
<td><strong>Behavioral factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical activity level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours of weekly activity, mean (SD)</td>
<td>10.5 (6.4)</td>
<td>10.0 (7.2)</td>
<td>0.96</td>
<td>.86</td>
</tr>
<tr>
<td>Eating/weight control behaviors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healthy weight control, mean (SD)</td>
<td>1.8 (1.6)</td>
<td>2.4 (1.5)</td>
<td>1.85</td>
<td>.03</td>
</tr>
<tr>
<td>Unhealthy weight control, mean (SD)</td>
<td>0.6 (1.1)</td>
<td>1.7 (1.7)</td>
<td>1.31</td>
<td>.18</td>
</tr>
<tr>
<td>Trying to gain weight, % (n)</td>
<td>26.6 (232)c</td>
<td>35.4 (5)c</td>
<td>1.77</td>
<td>.34</td>
</tr>
<tr>
<td>Trying to lose weight, % (n)</td>
<td>23.2 (192)c</td>
<td>42.1 (6)c</td>
<td>1.02</td>
<td>.98</td>
</tr>
<tr>
<td>Eating disorder diagnosis, % (n)</td>
<td>2.1 (22)</td>
<td>8.7 (2)</td>
<td>4.87</td>
<td>.08</td>
</tr>
<tr>
<td>Binge eating, % (n)</td>
<td>2.8 (29)</td>
<td>6.6 (1)</td>
<td>3.57</td>
<td>.19</td>
</tr>
<tr>
<td><strong>Substance use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cigarettes, % (n)</td>
<td>28.3 (284)</td>
<td>11.3 (2)</td>
<td>0.16</td>
<td>.09</td>
</tr>
<tr>
<td>Alcohol, % (n)</td>
<td>41.5 (416)</td>
<td>29.7 (5)</td>
<td>0.48</td>
<td>.23</td>
</tr>
<tr>
<td>Marijuana, % (n)</td>
<td>24.0 (239)</td>
<td>16.5 (3)</td>
<td>0.42</td>
<td>.29</td>
</tr>
<tr>
<td>Other, % (n)</td>
<td>6.3 (63)</td>
<td>5.4 (1)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

ORs and P values were derived from analyses adjusted for time 1 steroid use and propensity weights. ORs were standardized to allow comparison across scales. — indicates statistics were unable to be computed because of insufficient sample size.

* Percentage (n) values at time 1 of time 2 users and nonusers are reported for the dichotomous predictors; values for the continuous variables are means and SDs.

b Individuals tests may have different n values because of missing data on specific scales.

c The denominator of the proportion contains only the number of individuals answering “yes” to the question plus the number of individuals who answered that they have not tried to either gain or lose weight.
lences to be fairly stable or show only slight increases or declines after the turn of the century. In any case, our findings and those of others suggest that the press coverage in recent years surrounding steroid use by famous athletes and the congressional hearings into steroid use in baseball36 have not led to ever increasing levels of steroid use among adolescents.

The longitudinal nature of this investigation allowed us to properly examine developmental changes in steroid use as youth transition from middle school to high school, and from high school to young adulthood. Across the 5 years of follow-up, there were significant longitudinal reductions in use for older and younger boys and younger girls. There has been some cross-sectional support from other studies for decreased use across time. For instance, the YRBS in 2005 found that prevalences were higher for 9th-graders than for 12th-graders.7 The Monitoring the Future study reported lower prevalences for young adults compared with high school students, although they also found that older high school students had higher prevalences compared with younger adolescents. Our study, however, has the benefit of document-

<table>
<thead>
<tr>
<th>TABLE 4 Predictors of Time 2 Steroid Use in Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time 1 Variable</td>
</tr>
<tr>
<td>Personal factors</td>
</tr>
<tr>
<td>Weight and shape concerns</td>
</tr>
<tr>
<td>Weight Concerns, mean (SD)</td>
</tr>
<tr>
<td>Weight Importance, mean (SD)</td>
</tr>
<tr>
<td>Satisfaction, body build, mean (SD)</td>
</tr>
<tr>
<td>Satisfaction, shoulders, mean (SD)</td>
</tr>
<tr>
<td>Satisfaction, weight, mean (SD)</td>
</tr>
<tr>
<td>Ideal/estimated weight, mean (SD)</td>
</tr>
<tr>
<td>BMI, mean (SD)</td>
</tr>
<tr>
<td>Psychological variables</td>
</tr>
<tr>
<td>Self-esteem, mean (SD)</td>
</tr>
<tr>
<td>Depressed mood, mean (SD)</td>
</tr>
<tr>
<td>Suicide, thoughts, % (n)</td>
</tr>
<tr>
<td>Suicide, attempt, % (n)</td>
</tr>
<tr>
<td>Health/nutrition attitudes</td>
</tr>
<tr>
<td>Knowledge, healthy eating, mean (SD)</td>
</tr>
<tr>
<td>Efficacy, healthy food choices, mean (SD)</td>
</tr>
<tr>
<td>Concern about health, mean (SD)</td>
</tr>
<tr>
<td>Socio-environmental factors</td>
</tr>
<tr>
<td>Sports involvement</td>
</tr>
<tr>
<td>Weight-related sports, % (n)</td>
</tr>
<tr>
<td>Physical activity level</td>
</tr>
<tr>
<td>Hours of weekly activity, mean (SD)</td>
</tr>
<tr>
<td>Eating/weight control behaviors</td>
</tr>
<tr>
<td>Healthy weight control, mean (SD)</td>
</tr>
<tr>
<td>Unhealthy weight control, mean (SD)</td>
</tr>
<tr>
<td>Trying to gain weight, % (n)</td>
</tr>
<tr>
<td>Trying to lose weight, % (n)</td>
</tr>
<tr>
<td>Eating disorder diagnosis, % (n)</td>
</tr>
<tr>
<td>Binge eating, % (n)</td>
</tr>
<tr>
<td>Substance use</td>
</tr>
<tr>
<td>Cigarettes, % (n)</td>
</tr>
<tr>
<td>Alcohol, % (n)</td>
</tr>
<tr>
<td>Marijuana, % (n)</td>
</tr>
<tr>
<td>Other, % (n)</td>
</tr>
</tbody>
</table>

ORs and P-values were derived from analyses adjusted for time 1 steroid use and propensity weights. ORs were standardized to allow comparison across scales.

a Percentage (n) values at time 1 of time 2 users and nonusers are reported for the dichotomous predictors; values for the continuous variables are means and SDs.

b Individuals tests may have different n values because of missing data on specific scales.

c The denominator of the proportion contains only the number of individuals answering “yes” to the question plus the number of individuals who answered that they have not tried to either gain or lose weight.
ing decreased use across time in the same group of participants, confirming that use does in fact seem to decrease as youth progress through adolescence.

Because of the study design, we were also able to examine longitudinal predictors of steroid use over 5 years. Those boys who desired to weigh more and who used healthier weight-control behaviors were more likely to be steroid users after 5 years. The findings regarding healthy weight-control behaviors may reflect a general tendency to attend to eating and weight control, perhaps not yet developed to the point of being unhealthy, which may place boys at risk for engaging in steroid use later on. The lack of association between weight-related sports participation and steroid use in boys differs from the findings of other studies, including the longitudinal study by Dodge and Jaccard.11 We suspect it may be attributable to low power, because participation in weight-related sports at time 1 was nearly twice as high among time 2 steroid users as among nonusers. Another possible explanation is that boys' use is a risk factor for steroid use, but one that operates over a shorter period of time than the 5 years in our study. A better test of the influence of weight-related sports participation on steroid use might require a shorter time lag between participation and steroid use.

Steroid use in girls showed a marginally significant relationship with weight-related sports participation, with well over twice as many time 2 female steroid users as nonusers involved in weight-related sports at baseline. This finding begs additional exploration to determine whether and how strong a risk factor weight-related sports participation is for girls. Although the prevalence of steroid use is still low even among girls involved in weight-related sports, the identification of additional risk factors may allow us to define a high-risk group that could be targeted for prevention efforts.

Female adolescents who reported time 2 steroid use additionally had higher BMIs and were less satisfied with their weight at time 1. Thus, for girls, steroids may be a strategy to become leaner and more “toned,” whereas boys’ use is more often associated with wanting to increase their muscle mass and size. Interestingly, female time 2 steroid users were also more satisfied with their shoulders than time 2 nonusers. Female steroid users also had much lower levels of healthy nutrition knowledge and were less concerned about their health at time 1. This finding may be useful for identifying female adolescents at higher risk of steroid use in the future. One possible interpretation of these results is that female steroid users are typical of female adolescents in their desire to be smaller, but may have less concern for the health effects of the strategies they use to obtain their desired size.

The null findings regarding the use of other substances were unexpected. A number of studies have reported significant cross-sectional associations between the use of steroids and the use of other substances.14 However, in our study there were no significant associations, and in fact, for all of the substances studied, the prevalence of time 1 use of other substances was lower among time 2 steroid users than among nonusers. While keeping in mind the possibility of low power, this finding might be interpreted as indicating that drug use does not independently predict future steroid use above and beyond baseline levels.

Strengths of our study include its longitudinal design, its large, ethnically and socioeconomically diverse sample, and the many constructs assessed. However, there are limitations to our study that are important to consider. Despite the large sample size of our study, the relatively low, albeit disturbing, prevalence of steroid use indicates a need for interpreting our findings cautiously. Future research on steroid use should incorporate a larger sample size and more strategies to reduce attrition to capture greater numbers of steroid users. In addition, because steroid use was assessed with a single item regarding past year use, we were unable to determine the pattern of use during the 5-year study period. It would be of interest in future studies to be able to track steroid use over multiple, shorter intervals. Also, steroid use has been reported to frequently occur in monthly cycles, and questions regarding use should be designed to capture this cyclical pattern. Furthermore, given the proliferation of performance- and physique-enhancing substances available currently, there is always a concern that questions regarding the use of “steroids” may be misinterpreted by adolescents. Although it was not within the scope of our current study, future researchers might ask specifically about these substances (for instance, androstenedione or creatine) as well. In addition, it must be noted that attrition among time 1 steroid users was greater than among the other participants, so that associations may be biased if those steroid users who did not respond to EAT 2 were different from the ones who did respond, above and beyond what the propensity weighting could control.

Our study provides important information about steroid use in adolescence. These results are both encouraging, in that steroid use by participants declined as they became older, and concerning, in that over 1 in 100 adolescents (and in some groups much higher numbers) reported having used steroids at least once in the previous year. The narrowing gap between boys and girls observed in our investigation warrants additional study. Also, the fact that peak use occurred at younger ages and declined thereafter points to the usefulness of early prevention efforts, perhaps beginning in or even before the middle school years. Overall, the results of this study confirm the findings of others that steroid use is still a public health problem among adolescents, provide additional understanding of the natural course of steroid use.
over time, and offer insight into possible risk factors for future steroid use.

ACKNOWLEDGMENTS
This study was supported by grant R40 MC 00319 from the Maternal and Child Health Bureau (Title V, Social Security Act), Health Resources and Services Administration, Department of Health and Human Services. It was also supported, in part, by Adolescent Health Protection Research Training grant T01-DP000112 from the Centers for Disease Control and Prevention, Department of Health and Human Services.

REFERENCES
36. Curry J. The steroids hearings: baseball’s leaders; congress fires questions hard and inside, and baseball can only swing and miss. New York Times. March 18, 2005
Patient and Hospital Correlates of Clinical Outcomes and Resource Utilization in Severe Pediatric Sepsis

Folafoluwa O. Odetola, MD, MPH\textsuperscript{a,b}, Achamyeleh Gebremariam, MS\textsuperscript{b}, Gary L. Freed, MD, MPH\textsuperscript{b}

\textsuperscript{a}Division of Pediatric Critical Care Medicine, \textsuperscript{b}Department of Pediatrics and Communicable Diseases, Child Health Evaluation and Research Unit, University of Michigan, Ann Arbor, Michigan

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

ABSTRACT

OBJECTIVE. Our goal was to describe patient and hospital characteristics associated with in-hospital mortality, length of stay, and charges for critically ill children with severe sepsis.

METHODS. Our study consisted of a retrospective study of children 0 to 19 years of age hospitalized with severe sepsis using the 2003 Kids’ Inpatient Database. We generated national estimates of rates of hospitalization and then compared in-hospital mortality, length of stay, and total charges according to patient and hospital characteristics using multivariable regression methods. Severity of illness was measured by using all-patient refined diagnosis-related group severity of illness classification into minor, moderate, major, and extreme severity.

RESULTS. There were an estimated 21,448 hospitalizations for severe pediatric sepsis nationally in 2003. The in-hospital mortality rate was 4.2%. Comorbid illness was present in 34% of hospitalized children. Most (70%) of the extremely ill children were admitted to children’s hospitals. Length of stay was longer among patients with higher illness severity and nonsurvivors compared with survivors (13.5 vs 8.5 days). Hospitalizations at urban or children’s hospitals were also associated with longer length of stay than nonchildren’s or rural hospitals, respectively. Higher charges were associated with higher illness severity, and nonsurvivors had 2.5-fold higher total charges than survivors. Also, higher charges were observed among hospitalizations in urban or children’s hospitals. In multivariable regression analysis, multiple comorbid illnesses, multiple organ dysfunction, and greater severity of illness were associated with higher odds of mortality and longer length of stay. Higher hospital charges and longer length of stay were observed among transfer hospitalizations and among hospitalizations to children’s hospitals and nonchildren’s teaching hospitals compared with hospitals, which had neither children’s nor teaching status.

CONCLUSIONS. Mortality from severe pediatric sepsis is associated with patient illness severity, comorbid illness, and multiple organ dysfunction. Many characteristics are associated with resource consumption, including type of hospital, source of admission, and illness severity.
Pediatric sepsis that is associated with organ dysfunction, termed severe sepsis, is associated with higher mortality compared with sepsis without organ dysfunction.¹ Severe sepsis is a major cause of child mortality and morbidity in the United States and is also associated with high annual costs of $2 billion dollars.²,³ Although organ dysfunction progresses over time in the disease state, both in terms of severity and the number of organs involved, timely resuscitation of septic shock has been associated with improved outcomes assessed by reversal of organ dysfunction and improved survival.³

Although extensive research and clinical efforts are continually being made to reduce the morbidity and mortality associated with severe pediatric sepsis, no published information exists about potential variation that might occur in outcomes and resource utilization across different acute care settings. Also, the clinical course, sequelae, and outcomes of severe pediatric sepsis may be significantly impacted by the performance of time-sensitive and resource-dependent maneuvers, such as fluid resuscitation, interhospital transfer, subspecialist referral, and intensive care. The availability of, and access to, such key interventions might vary among different care settings, thereby providing an opportunity to investigate the association between various hospital characteristics and the outcomes and resource utilization for severe pediatric sepsis, while accounting for patient characteristics previously associated with such outcomes. Such information may be useful to clinicians and health care policy makers and highlight opportunities to alleviate illness burden.

Our study was conducted to describe hospital and patient characteristics that are independently associated with hospital mortality, length of stay (LOS), and charges for children hospitalized with severe sepsis, and to test the hypothesis that these measures would vary among different types of hospitals.

**METHODS**

**Study Design**

We conducted a retrospective study of hospitalized children with sepsis, 0 to 19 years old, who also had organ dysfunction involving at least 1 organ system. Our data source was the 2003 Kids’ Inpatient Database (KID). The KID was developed by the Agency for Healthcare Research and Quality (AHRQ) and includes nearly 3 million pediatric discharge records obtained from 3438 hospitals in 36 states. The KID is the only national, all-payer database of hospitalizations for children. The database contains 80% of the normal nonnewborn discharges from these states and is nationally representative with the inclusion of discharge weights in analyses. Information on patient demographics, hospital characteristics, and diagnosis codes is included for each hospitalization.

**Study Sample and Variable Identification**

Children with a primary or secondary diagnosis of sepsis associated with organ dysfunction were identified by using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes applying methodology described in previous studies.²,⁷ Severe sepsis was subsequently identified as the occurrence of sepsis and dysfunction of ≥1 organ system. Organ dysfunction was measured both according to the actual organ systems involved and the number of organ systems involved, because previous literature suggest worse outcomes with an increase in the number of organ systems involved in the disease state.¹,²,³ Patient characteristics other than organ dysfunction were age, gender, presence, and count of comorbid illness, and severity of illness using the all-patient refined diagnosis-related group (APRDRG) classification within the KID. The categories of APRDRG severity of illness were minor, moderate, major, and extreme loss of function, hereafter referred to as illness severity. The APRDRG severity classification is a proprietary, validated, and extensively used measure of illness severity that uses patient discharge data including principal and secondary diagnoses, procedures, and demographic information to assign patients to subclasses of illness severity. Because of small sample size for hospitalizations with minor illness severity, the categories for minor and moderate illness severity were combined for the analyses. Comorbidities were identified by using ICD-9-CM codes, applying methodology described in the literature. Comorbid illness was incorporated into the analyses because previous studies have associated the presence of comorbid illness with mortality² and cost of hospitalization.¹¹

Hospital characteristics were total bed size (large, medium, and small) defined by hospital location and teaching status, the type of hospital (children’s hospital status and teaching status: children’s teaching, children’s nonteaching, nonchildren’s teaching, and nonchildren’s nonteaching hospitals), source of admission (emergency department [ED], clinic, other hospitals), and location (urban or rural) as reported by publications available from AHRQ.⁸ Children’s hospitals were identified by the American Hospital Association Annual Survey of Hospitals and information from the National Association of Children’s Hospitals and Related Institutions. Teaching hospital status was determined by whether a hospital had an American Medical Association-approved residency program, was a member of the Council of Teaching Hospitals, or had a ratio of full-time equivalent interns and residents to beds of 0.25 or higher. Because of small sample size for admissions (77 admissions) to the nonteaching children’s hospitals, children’s teaching and nonteaching hospitals were combined for the analyses. The institutional review board of the University of Michigan Medical School approved the study.
Statistical Analysis

To identify factors associated with in-hospital mortality, LOS, and total charges, we initially identified the number of hospitalizations for severe sepsis in the KID and, subsequently, the frequency distribution of these discharges was obtained. Thereafter, rates of in-hospital mortality, and the frequency distribution of LOS and charges (with associated 95% confidence intervals [CIs]) for the hospitalizations was described according to both patient and hospital characteristics. The number of hospitalizations in our results was unweighted, although all effect estimates and accompanying 95% CIs were calculated by using sample weights to account for the complex survey design.

All estimates used the survey commands in Stata for Windows (Stata Corp, College Station, TX), which accounted for the complex survey design. To allow generation of stable estimates, cell frequencies that were too small (<70) for precise estimation were suppressed in the report as recommended by AHRQ. In these analyses, multivariable logistic regression, negative binomial regression, and multiple linear regression models for complex survey data were fit to assess differences in mortality, LOS, and total charges, respectively. Variance estimates accounted for clustering of data at the hospital level by using hospital identifiers as the primary sampling units.

RESULTS

There were 12,604 hospitalizations for severe sepsis in the database, representing 21,448 hospitalizations nationally in 2003. Over half (53%) of all hospitalized patients were male, and 58% were infants and children <5 years of age. Another peak in the frequency of hospitalizations (18%) was in children 15 to 19 years of age. Over half (59%) of hospitalizations were related to major or extreme illness severity (Table 1). For 58% of the hospitalizations, admission occurred via the ED. Approximately half (51%) of all cases were admitted into large-sized hospitals. The majority (98%) of all hospitalizations were in urban hospitals, and half of all hospitalizations were to children’s (teaching and nonteaching) hospitals.

The overall in-hospital mortality rate was 4.2%, and comorbid illness was present in 34% of all hospitalized children. Cardiac dysfunction occurred in 46.4% of the hospitalized children. Other dysfunctional organ systems were, in decreasing order of frequency: respiratory

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. (%) of Hospitalizations</th>
<th>% of all Deaths</th>
<th>% Mortality</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>3017 (24.0)</td>
<td>22.9</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>1–4</td>
<td>4214 (33.7)</td>
<td>25.9</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>5–9</td>
<td>1588 (12.7)</td>
<td>10.6</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>10–14</td>
<td>1424 (11.4)</td>
<td>16.2</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>15–19</td>
<td>2358 (18.3)</td>
<td>24.4</td>
<td>5.6</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>6663 (52.9)</td>
<td>52.8</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>5914 (47.1)</td>
<td>47.2</td>
<td>4.2</td>
<td>.95</td>
</tr>
<tr>
<td>Source of admission</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED</td>
<td>7219 (57.7)</td>
<td>47.0</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>Another hospital</td>
<td>1526 (12.1)</td>
<td>23.0</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Routine, clinic</td>
<td>3783 (30.2)</td>
<td>30.0</td>
<td>4.2</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>4253 (34.1)</td>
<td>63.4</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>8348 (65.9)</td>
<td>36.6</td>
<td>2.3</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Hospital bed size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>1843 (16.7)</td>
<td>16.8</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>3782 (32.5)</td>
<td>30.6</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>Large</td>
<td>6460 (50.8)</td>
<td>52.6</td>
<td>4.3</td>
<td>.81</td>
</tr>
<tr>
<td>Hospital location</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>1096 (10.2)</td>
<td>2.5</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>10,989 (89.8)</td>
<td>97.5</td>
<td>4.5</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Type of hospital</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children’s teaching and nonteaching hospital</td>
<td>5800 (51.4)</td>
<td>72.4</td>
<td>7.3</td>
<td></td>
</tr>
<tr>
<td>Nonchildren’s teaching hospital</td>
<td>2323 (18.6)</td>
<td>17.9</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>Nonchildren’s nonteaching hospital</td>
<td>3691 (30.9)</td>
<td>9.7</td>
<td>1.4</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>APRDRG levels of illness severity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor/moderate</td>
<td>5240 (41.3)</td>
<td>1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major</td>
<td>3708 (29.5)</td>
<td>14.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extreme</td>
<td>3653 (29.2)</td>
<td>84.4</td>
<td></td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>
(39.2%), hematologic (19.5%), neurologic (8.1%), renal (7%), and hepatic (0.7%). The overall mean LOS was 8.7 days (95% CI: 8.1–9.3 days), and the mean total charge per hospitalization was $47,126 (95% CI: $42,159–$52,097).

In bivariate analysis, fatal outcomes were associated with specific patient and hospital characteristics (Table 1). Nearly half of all fatalities occurred among children <5 years of age, and 25% occurred among children 15 to 19 years of age. There was no gender difference in mortality; however, patients with comorbid illness or extreme illness severity had higher mortality rates. Similarly, higher mortality rates were observed among hospitalizations via the ED, to an urban hospital, or to children’s hospitals.

Analysis of illness severity revealed that 40% of hospitalizations to children’s hospitals had extreme illness severity, much higher than in nonchildren’s hospitals with teaching or nonteaching status (Table 2). Likewise, 70% of all hospitalized children with extreme illness severity were admitted into children’s hospitals compared with 18% and 12% to nonchildren’s teaching and nonchildren’s nonteaching hospitals, respectively. Of note, half of all transfer hospitalized children had extreme illness severity, much higher than for those with any other source of admission (Table 2).

LOS was 5 days longer among nonsurvivors and was notably longer among patients with major or extreme illness severity compared with minor/moderate severity (Table 3). Hospitalizations into urban hospitals or children’s hospitals were associated with longer LOS compared with rural or nonchildren’s hospitals.

Total hospital charges were 2.5-fold higher among nonsurvivors when compared with survivors (Table 4). Higher charges were also associated with higher illness severity, and among hospitalizations into urban hospitals or children’s hospitals, compared with rural or nonchildren’s hospitals, respectively.

Multivariable logistic regression analysis revealed that the number of comorbid illnesses (odds ratio [OR]: 1.54; 95% CI: 1.38–1.73), number of dysfunctional organ systems (OR: 2.25; 95% CI: 2.02–2.51), and severity of illness were associated with mortality (Table 5). There was no statistically significant difference in risk-adjusted mortality according to the type of hospital (children’s/teaching hospital status), despite hospitalization of children with the highest illness severity within children’s hospitals.

In multivariable negative binomial regression analysis, longer LOS was associated with multiple comorbid illnesses, multiple organ dysfunction, and higher illness severity (Table 6). Hospitalizations from either the clinic or as transfer from other hospitals were associated with higher LOS compared with ED admissions. Hospitalization into children’s hospitals or nonchildren’s teaching hospitals had longer LOS compared with hospitals with neither children’s nor teaching status, after controlling for patient factors.

Hospital charges for children with severe sepsis varied by patient and hospital characteristics. Multivariable lin-

### Table 2: Illness Severity According to Patient and Hospital Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minor/Moderate</th>
<th>Major</th>
<th>Extreme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>1506 (49.6)</td>
<td>689 (22.9)</td>
<td>822 (27.5)</td>
</tr>
<tr>
<td>1–4</td>
<td>2320 (54.6)</td>
<td>1021 (24.4)</td>
<td>875 (21.0)</td>
</tr>
<tr>
<td>5–9</td>
<td>519 (32.3)</td>
<td>573 (36.4)</td>
<td>496 (31.3)</td>
</tr>
<tr>
<td>10–14</td>
<td>310 (21.4)</td>
<td>540 (37.8)</td>
<td>575 (40.8)</td>
</tr>
<tr>
<td>15–19</td>
<td>585 (24.6)</td>
<td>885 (37.4)</td>
<td>888 (38.0)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>2768 (41.2)</td>
<td>1889 (28.5)</td>
<td>2009 (30.4)</td>
</tr>
<tr>
<td>Female</td>
<td>2451 (41.5)</td>
<td>1819 (30.6)</td>
<td>1644 (27.9)</td>
</tr>
<tr>
<td>Source of admission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED</td>
<td>3323 (45.8)</td>
<td>2120 (29.4)</td>
<td>1778 (24.8)</td>
</tr>
<tr>
<td>Another hospital</td>
<td>323 (20.8)</td>
<td>435 (28.5)</td>
<td>769 (50.7)</td>
</tr>
<tr>
<td>Routine, clinic</td>
<td>1562 (41.0)</td>
<td>1134 (30.1)</td>
<td>1087 (28.9)</td>
</tr>
<tr>
<td>Hospital bed size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>698 (37.8)</td>
<td>589 (31.7)</td>
<td>533 (29.3)</td>
</tr>
<tr>
<td>Medium</td>
<td>1622 (42.1)</td>
<td>1047 (27.8)</td>
<td>1050 (28.5)</td>
</tr>
<tr>
<td>Large</td>
<td>2562 (39.8)</td>
<td>1916 (29.7)</td>
<td>1883 (29.0)</td>
</tr>
<tr>
<td>Hospital location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>712 (64.7)</td>
<td>258 (23.4)</td>
<td>100 (9.4)</td>
</tr>
<tr>
<td>Urban</td>
<td>4170 (37.4)</td>
<td>3294 (30.1)</td>
<td>3366 (31.1)</td>
</tr>
<tr>
<td>Type of hospital</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children’s teaching and nonteaching hospital</td>
<td>1635 (28.5)</td>
<td>1860 (32.0)</td>
<td>2307 (39.5)</td>
</tr>
<tr>
<td>Nonchildren’s teaching hospital</td>
<td>988 (42.4)</td>
<td>686 (29.7)</td>
<td>649 (28.0)</td>
</tr>
<tr>
<td>Nonchildren’s nonteaching hospital</td>
<td>2326 (63.7)</td>
<td>921 (24.7)</td>
<td>445 (11.6)</td>
</tr>
</tbody>
</table>
ear regression analysis revealed higher charges among patients who were admitted on transfer from other hospitals (Table 7). Higher charges were also observed among hospitalizations in children’s hospitals and nonchildren’s teaching hospitals compared with hospitals with neither children’s nor teaching status. Likewise, urban hospitals had higher charges compared with rural hospitals. Higher charges were also observed with increasing illness severity, multiple comorbid illnesses or multiple organ dysfunction.

### DISCUSSION

This is the first study, to our knowledge, to describe an association between characteristics of the hospitals where care is provided to children with severe sepsis and in-hospital mortality, LOS, and charges. We found higher charges and longer length of stay for children admitted to children’s hospitals and urban hospitals compared with admission to nonchildren’s or rural hospitals, respectively. Of note, although most of the children with extreme illness severity were admitted to children’s hospitals, there was no statistically significant difference in risk-adjusted mortality between these hospitalizations and those within nonchildren’s hospitals. This study also corroborates previous studies that have described a significant association between the presence of comorbid illness and multiple organ dysfunction with mortality and LOS for children hospitalized with severe sepsis.

#### TABLE 3: LOS Associated With Hospitalizations for Severe Pediatric Sepsis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean LOS, d</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Survival status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survived</td>
<td>8.5</td>
<td>7.9–9.1</td>
</tr>
<tr>
<td>Died</td>
<td>13.5</td>
<td>10.8–16.2</td>
</tr>
<tr>
<td><strong>Hospital bed size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>9.3</td>
<td>8.1–10.4</td>
</tr>
<tr>
<td>Medium</td>
<td>8.6</td>
<td>7.0–10.1</td>
</tr>
<tr>
<td>Large</td>
<td>8.4</td>
<td>7.8–9.0</td>
</tr>
<tr>
<td><strong>Hospital location</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>3.4</td>
<td>2.9–3.8</td>
</tr>
<tr>
<td>Urban</td>
<td>9.2</td>
<td>8.5–9.9</td>
</tr>
<tr>
<td><strong>Type of hospital</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children’s teaching and nonteaching hospitals</td>
<td>11.4</td>
<td>10.2–12.6</td>
</tr>
<tr>
<td>Nonchildren’s teaching hospital</td>
<td>8.0</td>
<td>7.3–8.7</td>
</tr>
<tr>
<td>Nonchildren’s nonteaching hospital</td>
<td>4.3</td>
<td>4.0–4.6</td>
</tr>
<tr>
<td><strong>APRDRG levels of illness severity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor/moderate</td>
<td>3.26</td>
<td>3.09–3.43</td>
</tr>
<tr>
<td>Major</td>
<td>7.90</td>
<td>7.49–8.31</td>
</tr>
<tr>
<td>Extreme</td>
<td>17.31</td>
<td>16.47–18.14</td>
</tr>
</tbody>
</table>

* Statistically significant because 95% CIs do not overlap.

#### TABLE 4: Charges Associated With Hospitalizations for Severe Pediatric Sepsis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean Charges, $</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Survival status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survived</td>
<td>44 270.7</td>
<td>39 628.2–48 913.2</td>
</tr>
<tr>
<td>Died</td>
<td>113 431.5</td>
<td>94 667–132 195</td>
</tr>
<tr>
<td><strong>Hospital bed size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>47 335.6</td>
<td>39 979.6–54 691.5</td>
</tr>
<tr>
<td>Medium</td>
<td>50 686.1</td>
<td>36 383.0–64 989.2</td>
</tr>
<tr>
<td>Large</td>
<td>44 070.8</td>
<td>39 645.5–48 496.1</td>
</tr>
<tr>
<td><strong>Hospital location</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>9099.4</td>
<td>7376.2–10 822.6</td>
</tr>
<tr>
<td>Urban</td>
<td>51 088.2</td>
<td>45 341.3–56 834.9</td>
</tr>
<tr>
<td><strong>Type of hospital</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children’s teaching and nonteaching hospitals</td>
<td>65 428.5</td>
<td>55 075.3–75 781.7</td>
</tr>
<tr>
<td>Nonchildren’s teaching hospital</td>
<td>42 820.5</td>
<td>36 493.7–49 147.3</td>
</tr>
<tr>
<td>Nonchildren’s nonteaching hospital</td>
<td>18 534.7</td>
<td>16 159.6–20 909.8</td>
</tr>
<tr>
<td><strong>APRDRG levels of illness severity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor/moderate</td>
<td>11 373.0</td>
<td>10 402.5–12 343.5</td>
</tr>
<tr>
<td>Major</td>
<td>36 717.4</td>
<td>33 286.5–40 148.4</td>
</tr>
<tr>
<td>Extreme</td>
<td>108 657.5</td>
<td>99 787.6–117 527.3</td>
</tr>
</tbody>
</table>

* Statistically significant because 95% CIs do not overlap.

#### TABLE 5: Patient and Hospital Characteristics Associated With In-hospital Mortality in a Multivariable Logistic Regression Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Adjusted OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>1.00</td>
<td>0.99–1.02</td>
<td>.76</td>
</tr>
<tr>
<td>Female vs male gender</td>
<td>0.99</td>
<td>0.82–1.22</td>
<td>.97</td>
</tr>
<tr>
<td>Per comorbid illness</td>
<td>1.55</td>
<td>1.38–1.73</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Per dysfunctional organ system</td>
<td>2.25</td>
<td>2.02–2.51</td>
<td>&lt;.01</td>
</tr>
<tr>
<td><strong>APRDRG levels of illness severity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor or moderate</td>
<td>0.03</td>
<td>0.01–0.07</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Major</td>
<td>0.29</td>
<td>0.22–0.37</td>
<td>&lt;.01</td>
</tr>
<tr>
<td><strong>Source of admission</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Another hospital</td>
<td>1.27</td>
<td>0.99–1.61</td>
<td>.05</td>
</tr>
<tr>
<td>Routine/clinic</td>
<td>1.17</td>
<td>0.94–1.45</td>
<td>.16</td>
</tr>
<tr>
<td><strong>Type of hospital</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children’s teaching and nonteaching hospital</td>
<td>1.48</td>
<td>1.00–2.19</td>
<td>.05</td>
</tr>
<tr>
<td>Nonchildren’s teaching hospital</td>
<td>1.40</td>
<td>0.91–2.15</td>
<td>.12</td>
</tr>
<tr>
<td><strong>Hospital bed size</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>1.00</td>
<td>0.70–1.45</td>
<td>.98</td>
</tr>
<tr>
<td>Large</td>
<td>1.14</td>
<td>0.83–1.57</td>
<td>.43</td>
</tr>
<tr>
<td><strong>Hospital location</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban vs rural</td>
<td>1.14</td>
<td>0.60–2.19</td>
<td>.69</td>
</tr>
</tbody>
</table>

* Age was analyzed as a continuous variable. The adjusted OR reflects the odds of the specified outcome compared with the reference category within the same variable, adjusted for the other outcome categories that appear here.

* Reference category was extreme illness severity.

* Reference category was the ED.

* Reference category was nonchildren’s nonteaching hospital.

* Reference category was small hospital bed size.
orates previous findings among critically ill adults that higher resource utilization occurred among certain patients who underwent interhospital transfer.12,13

Children’s hospitals, regardless of their teaching status, along with nonchildren’s teaching hospitals, had significantly longer LOS of hospitalizations when compared with hospitals without teaching or children’s hospital status. A previous study highlighted variation in LOS by hospital type among hospitalizations for common pediatric conditions, without clearly defined reasons for such observation.14 A significant limitation of that study was the inability to adjust for illness severity. However, in our study, similar findings were observed after rigorous adjustment for illness severity by using available methods. We speculate that this differential LOS and resource consumption according to the type of hospital for severe sepsis hospitalizations might be attributable to extensive in-hospital rehabilitation and multidisciplinary care for the most severely ill patients selectively admitted to children’s hospitals. Outside illness severity, however, multiple patient characteristics, practice patterns, or institutional factors may cause the wide differences across hospitals in LOS.15 Additional study of differences in LOS among patients with severe sepsis according to the type of hospital might help elucidate whether these or other factors are germane to the critically ill pediatric population with severe sepsis.

A previous study of hospitalizations for common pediatric conditions16 reported significant differences in charges by hospital type, similar to our study. It has been speculated that subspecialty care and use of advanced technology within children’s hospitals might explain some of the differential charges reported. Furthermore, previous studies have reported higher charges among adult hospitalizations in tertiary hospitals,17 suggesting elevated hospital charges might occur with specialization of medical care.

Fatal outcomes for children with severe sepsis were associated more strongly with patient characteristics than hospital characteristics. Historically, a significant limitation of the use of administrative databases for outcomes research pertained to the inability to control for patient severity of illness, leading to the use of various methods over time.18,19 In our study, however, available methodology was rigorously applied to account for comorbid illness and multiplicity of organ dysfunction as markers of illness severity. In addition, the use of APRDRG severity provided an opportunity for categorization of patients according to level of illness severity.

The finding of an association between multiple organ dysfunction and death implies an urgent need for concerted clinical and educational efforts within clinical care settings aimed at limiting progression of illness severity and organ dysfunction in children with severe sepsis. This finding indirectly corroborates previous studies1,4 that emphasize the need for early resuscitation to pre-

---

**TABLE 6**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Adjusted Rate Ratio</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female vs male gender</td>
<td>1.01</td>
<td>0.97–1.05</td>
<td>.59</td>
</tr>
<tr>
<td>Per comorbid illness</td>
<td>1.15</td>
<td>1.11–1.20</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Per dysfunctional organ system</td>
<td>1.17</td>
<td>1.12–1.22</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>APRDRG levels of illness severity&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor or moderate</td>
<td>0.27</td>
<td>0.24–0.30</td>
<td>.95</td>
</tr>
<tr>
<td>Major</td>
<td>0.54</td>
<td>0.50–0.57</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Source of admission&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Another hospital</td>
<td>1.28</td>
<td>1.19–1.38</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Routine/clinic</td>
<td>1.09</td>
<td>1.02–1.16</td>
<td>.01</td>
</tr>
<tr>
<td>Type of hospital&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children’s teaching and non teaching hospital</td>
<td>1.43</td>
<td>1.31–1.56</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Hospital nonchildren’s teaching hospital</td>
<td>1.25</td>
<td>1.16–1.34</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Hospital bed size&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>0.98</td>
<td>0.86–1.11</td>
<td>.73</td>
</tr>
<tr>
<td>Large</td>
<td>1.05</td>
<td>0.95–1.16</td>
<td>.32</td>
</tr>
<tr>
<td>Hospital location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban vs rural</td>
<td>1.28</td>
<td>1.20–1.38</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

The adjusted rate ratio reflects the odds of the specified outcome compared with the reference category within the same variable, adjusted for the other outcome categories that appear here.

<sup>a</sup> Reference category was extreme illness severity.
<sup>b</sup> Reference category was the emergency department.
<sup>c</sup> Reference category was nonchildren’s nonteaching hospital.
<sup>d</sup> Reference category was small hospital bed size.

**TABLE 7**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Adjusted Charge Ratio</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female vs male gender</td>
<td>1.00</td>
<td>0.97–1.04</td>
<td>.91</td>
</tr>
<tr>
<td>Per comorbid illness</td>
<td>1.18</td>
<td>1.13–1.23</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Per dysfunctional organ system</td>
<td>1.41</td>
<td>1.36–1.47</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Source of admission&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Another hospital</td>
<td>1.23</td>
<td>1.12–1.35</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Routine/clinic</td>
<td>0.95</td>
<td>0.87–1.03</td>
<td>.21</td>
</tr>
<tr>
<td>Type of hospital&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children’s teaching hospital</td>
<td>1.19</td>
<td>1.06–1.35</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Hospital location</td>
<td>1.49</td>
<td>1.32–1.70</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Hospital bed size&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>1.11</td>
<td>0.93–1.32</td>
<td>.24</td>
</tr>
<tr>
<td>Large</td>
<td>1.11</td>
<td>0.98–1.26</td>
<td>.11</td>
</tr>
<tr>
<td>Hospital location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban vs rural</td>
<td>1.69</td>
<td>1.54–1.86</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>APRDRG levels of illness severity&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor or moderate</td>
<td>0.18</td>
<td>0.17–0.20</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Major</td>
<td>0.44</td>
<td>0.41–0.46</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

The adjusted charge ratio reflects the odds of the specified outcome compared with the reference category within the same variable, adjusted for the other outcome categories that appear here.

<sup>a</sup> Reference category was the ED.
<sup>b</sup> Reference category was nonchildren’s nonteaching hospital.
<sup>c</sup> Reference category was small hospital bed size.
<sup>d</sup> Reference category was extreme illness severity.
vent or limit organ dysfunction as a major mortality risk-reduction strategy in severe pediatric sepsis.

Resource utilization, measured by LOS and hospital charges, was associated with illness severity and varied significantly by hospital characteristics, including type of hospital and geographical location. Efforts to alleviate this resource burden, particularly among children’s hospitals and hospitals located in urban regions, will need to develop better understanding of circumstances surrounding interhospital transfers, timing of resuscitative care, and other health system processes that might be incorporated into the care of children with severe sepsis.

The findings of our study should be interpreted in light of certain limitations. The KID is a database of administrative discharge data without clinical information beyond what can be captured in ICD-9-CM diagnosis and procedure codes. Therefore, it was not possible to study the clinical course for each patient, the need for and receipt of various therapeutic interventions, and how clinical care was coordinated at the patient and hospital level. An additional limitation is that the identification of cases of severe sepsis was ascertained via ICD-9-CM diagnosis codes and is susceptible to inaccuracies of detection and attribution that may have biased our findings.

Also, although the care of patients with severe sepsis might be improved by early resuscitation in the community before transfer to referral hospitals, determination of the characteristics of the referring hospitals and the circumstances surrounding transfer among the transfer hospitalizations was not possible in our study. This limited the ability to investigate any delay in resuscitative care, which might impact clinical outcomes and resource utilization at the receiving hospital. The characterization of types of hospitals within the database did not permit identification of which hospitals were freestanding children’s hospitals, which are likely to be highly specialized with extensive resource capability. The KID contains nonclinical administrative data collected during patient hospitalizations and has no follow-up or longitudinal information on patients after hospital discharge. This limitation did not permit determination of patient morbidity and functional status.

CONCLUSIONS
Mortality from severe pediatric sepsis was influenced largely by patient characteristics including comorbid illness and multiple organ system dysfunction. Resource consumption, assessed by LOS and hospital charges, was influenced both by these patient characteristics, and multiple hospital characteristics, such as type and location of hospital and source of admission.

Efforts to curb mortality from severe pediatric sepsis will need to address patient illness severity and explore mechanisms to stem the progression of multiorgan system dysfunction. The success of such mortality-reduction strategies and others aimed at alleviating both the overall and differential resource burden associated with severe pediatric sepsis will hinge on improved understanding of the role of patient triage, interhospital transfer, hospital characteristics, and other yet unmeasured factors on clinical outcomes and resource utilization within various acute care settings.

REFERENCES
WHAT $1.2 TRILLION CAN BUY

“The human mind isn’t very well equipped to make sense of a figure like $1.2 trillion. We don’t deal with a trillion of anything in our daily lives, and so when we come across such a big number, it is hard to distinguish it from any other big number. Millions, billions, a trillion—they all start to sound the same. The way to come to grips with $1.2 trillion is to forget about the number itself and think instead about what you could buy with the money. When you do that, a trillion stops sounding anything like millions or billions.

For starters, $1.2 trillion would pay for an unprecedented public health campaign—a doubling of cancer research funding, treatment for every American whose diabetes or heart disease is now going unmanaged and a global immunization campaign to save millions of children’s lives. Combined, the cost of running those programs for a decade wouldn’t use up even half our money pot. So we could then turn to poverty and education, starting with universal preschool for every 3- and 4-year-old child across the country. The city of New Orleans could also receive a huge increase in reconstruction funds. The final big chunk of the money could go to national security. The recommendations of the 9/11 Commission that have not been put in place—better baggage and cargo screening, stronger measures against nuclear proliferation—could be increased to beat back the Taliban’s recent gains, and a peacekeeping force could put a stop to the genocide in Darfur. All that would be one way to spend $1.2 trillion. Here would be another: The war in Iraq.”

Noted by JFL, MD
Risk of Recurrent Childhood Arterial Ischemic Stroke in a Population-Based Cohort: The Importance of Cerebrovascular Imaging

Heather J. Fullerton, MD, MASa,b, Yvonne W. Wu, MD, MPHa,b, Stephen Sidney, MD, MPHc, S. Claiborne Johnston, MD, MPH, PhDad

Departments of aNeurology, bPediatrics, and dEpidemiology, University of California, San Francisco, California; cDivision of Research, Kaiser Permanente Medical Care Program, Oakland, California

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

ABSTRACT

OBJECTIVE. Few data exist regarding rates and predictors of recurrence after childhood arterial ischemic stroke. We sought to establish such rates within a large, multi-ethnic population and determine whether clinical vascular imaging predicts recurrence.

PATIENTS AND METHODS. In a population-based cohort study, we collected data on all documented cases of arterial ischemic stroke among 2.3 million children (<20 years old) enrolled in a northern Californian managed care plan from January 1993 to December 2004. Perinatal strokes were those that occurred by 28 days of life. Data on cerebrovascular imaging (conventional or magnetic resonance angiography), including presence of vascular abnormalities, were abstracted from official radiology reports. We used Kaplan-Meier survival-analysis techniques to determine rates and predictors of recurrent stroke.

RESULTS. Among 181 incident childhood stroke cases (84 perinatal; 97 later childhood), there were 16 recurrent strokes (1 after a perinatal stroke) at a median of 2.7 months. The 5-year cumulative recurrence rates were 1.2% after perinatal stroke and 19% after later childhood stroke. Of the 97 children with later childhood strokes, 52 received cerebrovascular imaging, predominantly magnetic resonance angiography (n = 36) and conventional angiography (n = 26). Although there were no recurrences among children with normal vascular imaging, children with a vascular abnormality had a 5-year cumulative recurrence rate of 66%.

CONCLUSIONS. Strokes recur in one fifth of cases of later childhood arterial ischemic stroke but are rare after perinatal stroke. Among the later childhood cases, cerebrovascular imaging identifies those at highest risk for recurrence.
Despite increasing recognition of arterial ischemic stroke (AIS) as an important cause of childhood disability, few data exist regarding rates and predictors of recurrent childhood stroke. Although children with perinatal strokes are thought to be at lower risk than older children, a direct comparison of perinatal versus later childhood strokes within the same population has never been made. Stroke etiology also likely impacts recurrence risk. Conditions associated with first childhood stroke include congenital heart disease, sickle cell disease, meningitis, a variety of hereditary and acquired prothrombotic states, and a number of vasculopathies.

In a study of children identified in several German hospitals, children whose strokes were classified as “vascular” in etiology had fourfold odds of suffering a recurrence compared with children with idiopathic strokes, but this finding has not been validated in a distinct clinical setting or a more diverse ethnic population.

The dearth of information on recurrence after childhood stroke likely contributes to practice variability in the evaluation and treatment of childhood stroke. Several paragraphs embedded within the American College of Chest Physicians’ guidelines on antithrombotic therapy in children comprise the sole in-print guidelines addressing childhood AIS. There are no established standards of care for the diagnostic evaluation of children with stroke, and head and vascular imaging are used variably. Furthermore, secondary prevention practices vary widely, and many children receive no specific intervention to reduce risk of recurrence.

We sought to study the risk of stroke recurrence among unselected children and neonates in a defined multiethnic population. The objectives of this study were to determine (1) recurrence rates after childhood AIS during long-term follow-up, (2) the relative risk of recurrence in children with perinatal versus later childhood AIS, and (3) the utility of clinical cerebrovascular imaging in predicting recurrence risk.

Methods

After obtaining the approval of the Kaiser Permanente Medical Care Program (KPMCP) and University of California, San Francisco, institutional review boards, we performed a retrospective cohort study in which we sought to identify all strokes occurring within the population of children (<20 years of age) enrolled in KPMCP during an 11-year study period (January 1993 through December 2003). This report describes only the children with AIS.

Setting

KPMCP is the largest nonprofit managed care organization in the country, with 16 hospitals and 36 outpatient facilities in northern California. It provides typically long-term care for ~3.1 million people, or 30% of the regional population. The KPMCP population shares the general sociodemographic distribution of California except for an underrepresentation of socioeconomic extremes.

Case Ascertainment

Cases were ascertained through a multitiered process. First, we electronically searched the KPMCP patient databases for hospital discharge diagnoses (coded by a medical charts abstractor) and outpatient diagnoses (coded by the treating physician) suggestive of a stroke. This included diagnoses of ischemic stroke (International Classification of Diseases, Ninth Revision [ICD-9] codes 433–436, 437.6, and 325), subarachnoid hemorrhage (ICD-9 codes 430 and 772.2), and intracerebral hemorrhage (ICH; ICD-9 code 431). The search included all out-of-plan hospitalizations. Second, we performed keyword searches of all electronic head imaging reports (MRI and computed tomography scan) with the following text strings: stroke, infarct, infarction, infarcted, thrombus, thromboembolic, thromboembolism, thrombotic, thrombosis, ischemia, ischemic, lacune, lacunar, vascular event, porencephaly, and porencephalic. A single pediatric stroke neurologist (Dr Fullerton) reviewed all identified reports. Third, we cross-referenced with previous studies of cerebral palsy and perinatal arterial ischemic stroke that used a KPMCP birth cohort (January 1991 through December 2002).

Potential cases were subjected to chart review. Two child neurologists (Drs Fullerton and Wu) independently confirmed cases of childhood stroke; a third stroke neurologist (Dr Johnston) arbitrated disputes. The criteria for AIS were: (1) documented clinical presentation consistent with stroke, such as a sudden onset focal neurologic deficit; and (2) computed tomography scan or MRI showing a focal ischemic infarct in a location and of a maturity consistent with the neurologic signs and symptoms. Cases were excluded if the stroke occurred before the child’s enrollment in Kaiser or outside of the study period.

Perinatal strokes were those occurring between 28 weeks’ gestation and 28 days of life. They included presumed perinatal strokes, those that presumably occur perinatally but are not diagnosed until the children become symptomatic later in life. Strokes occurring after 28 days were called later childhood strokes.

Data Abstraction

A single pediatric nurse medical charts analyst reviewed records of all confirmed cases by using a standardized protocol. Dr Fullerton rereviewed all records and used all available data to categorize the stroke etiology: cardiac (including congenital and valvular heart disease), infection (including meningitis, encephalitis, and sepsis), hypercoaguable state (acquired or hereditary), primary vasculopathy (arterial dissection, moyamoya [a progressive occlusive vasculopathy of the intracranial blood ves-
sels], vasculitis, and arterial stenosis in the absence of another etiology, such as meningitis), other, or idiopathic. 2,8

Vascular Imaging
We abstracted data regarding any cerebral vascular imaging (magnetic resonance angiography, computed tomography angiography, or catheter angiography) by using the official interpretation by the clinical radiologist. We defined “abnormal vascular imaging” as any “stenosing” vascular abnormality: arterial stenosis, moyamoya, arterial dissection, fibromuscular dysplasia, or vasculitis.2,8 Our definition excluded isolated arterial occlusion, which is more likely to represent thrombus than a vascular abnormality.2

Recurrent Cerebrovascular Events
KPMCP members receive all routine care within the system, facilitating the acquisition of follow-up data. Out of plan services, including emergent admissions to outside hospitals, are recorded in an electronic database, and copies of medical charts from outside admissions are typically available in outpatient charts. We abstracted data regarding any additional strokes or transient ischemic attacks (TIAs) that occurred after the index AIS, applying the same adjudication procedures but blinding the adjudicators to the vascular imaging reports. Because children with certain ischemic stroke etiologies (eg, moyamoya) are also at risk for hemorrhagic stroke, hemorrhagic strokes after the index AIS were considered recurrences. Similar clinical and radiologic criteria were used for this stroke subtype. Hemorrhagic transformation of the initial infarct was not considered a recurrence. Criteria for the diagnosis of TIA included (1) a focal neurologic deficit of acute onset lasting <24 hours, (2) no radiographic evidence of an infarct, and (3) clinical suspicion of a TIA by a physician.

Data Analysis
Incidence rates were calculated as the number of strokes divided by the number of person-years at risk. We used survival analysis to determine recurrence rates. Our primary outcome variable was time to recurrent stroke; for this analysis, the period at risk began on the date of the index AIS and ended on the date of the first recurrent stroke (the “failure” event) or censoring. Cases were censored (ie, withdrawn from the survival analysis) at either death or loss to follow-up by using the date of the last recorded visit to a KPMCP facility. Our primary outcome variable was time to recurrent stroke defined as date of the index AIS to date of the first recurrent stroke. Cumulative recurrence rates were derived from hazard functions. A secondary analysis used the combined outcome of stroke or TIA.

We constructed Kaplan-Meier survival curves to compare recurrence-free survival rates between different subgroups and used log-rank tests to determine the significance of univariate associations (α set at .05). Univariate Cox proportional hazards regression techniques were used to assess relative risk in terms of hazard ratios (HRs) with 95% confidence intervals (CIs). We used log-minus-log survival plots to demonstrate proportionality for binary predictors.14 However, we primarily used techniques that are not dependent on the proportional hazards assumption: comparisons of cumulative recurrence rates and log-rank tests.

χ² tests were used to compare simple proportions. We used the nonparametric test for trends across ordered groups to determine whether the proportion of children receiving vascular imaging changed over time.17 We used Stata 9.0 (Stata Corp, College Station, TX) to perform all statistical calculations.

RESULTS
Our study of 8 963 308 person-years included 2 347 982 individual children enrolled in KPMCP for a mean of 3.8 years during the 11-year study period. A total of 181 cases of childhood AISs were confirmed, yielding an overall annual incidence of 2.0 per 100 000 person-years. Of these, 84 (46%) were perinatal; 97 (54%) occurred in later childhood. The vast majority (98%) were first-time strokes; only 3 children had previous strokes (before the study period or before enrollment in KPMCP). The stroke cohort was ethnically diverse (16% black, 46% non-Hispanic white, 19% Hispanic, 12% Asian, 1% Native American, and 6% other or unknown), with a preponderance of boys (55%). Etiologies of the index strokes are shown in Table 1. Antithrom-
botic agents were used after 2% of perinatal strokes and after 51% of later childhood strokes (Table 1).

A total of 64 children had cerebrovascular imaging; such imaging was uncommonly performed after perinatal stroke (Table 2). Compared with the 52 older children who received vascular imaging, the 46 who did not receive vascular imaging were more likely to have either a cardiac etiology (22% vs 4%; \(P = .006\)) or infectious etiology (36% vs 12%; \(P = .005\)). There was no change in rates of vascular imaging over the 11 years of study (\(P = .91\)).

Follow-up data were available for 84 (100%) children with perinatal strokes and 92 (95%) children with later childhood strokes. Only 6 children died during the acute period (2 neonates). The perinatal stroke group was followed for a median of 5.8 years (mean: 6.0 years [range: 3 days to 12.4 years]), compared with a median of 5.1 years (mean: 5.0 years [range: 5 days to 11.8 years]) for the later childhood stroke group.

Recurrent Events

We identified a total of 16 subjects with first recurrent strokes occurring at a median of 2.2 months after the index stroke (range: 1 day to 4 years). Fourteen recurrences were AIS and 2 were ICH; only 1 (an ICH) occurred after a perinatal stroke. Etiologies of the first recurrent events are shown in Table 1. Six recurrences (all AIS) were in children treated with antithrombotic agents (3 received aspirin, 2 low molecular weight heparin, and 1 warfarin), whereas 10 (including the 2 ICH) were in children who received no antithrombotic therapy (Table 1). Although hemorrhagic transformation of the initial AIS was not considered a recurrent event, there was no ICH among the children treated with antithrombotic agents.

Recurrence After Later Childhood Stroke

The cumulative stroke recurrence rate was 15% (95% CI: 12%–30%) at 1 year, and 19% (95% CI: 12%–30%) at 5 years (Fig 1). Eleven of the 15 recurrences occurred within the first 6 months, and only 2 occurred beyond 1 year. The single ICH after a later childhood stroke was in a child with systemic lupus erythematosus who subsequently had 2 recurrent AISs 4 months after the ICH. There were 4 additional children who had recurrent TIAs without recurrent strokes. The 5-year cumulative recurrence rate for any cerebrovascular event (stroke or TIA) was 24% (95% CI: 16%–36%). Of the 19 subjects with recurrent cerebrovascular events, 11 had multiple recurrences, with 6 having \(\geq 3\) recurrences.

Recurrence After Perinatal Stroke

The single recurrence after perinatal stroke was in a neonate with meningoencephalitis who suffered an ICH 7 days after a small internal capsule AIS. Overall, children with perinatal strokes were 94% less likely to suffer a recurrent stroke than those with later childhood strokes (HR: 0.06; 95% CI: 0.008–0.48; \(P = .008\)). Conversely, children with later childhood strokes had a 16-fold increased hazard of recurrence compared with the neonates (HR: 16; 95% CI: 2.1–120; \(P = .008\)).

Abnormal Vascular Imaging Predicts Recurrence

We limited this analysis to the higher-risk later childhood subgroup. Of the 15 recurrent strokes, 13 occurred among children who had received vascular imaging. Abnormal vascular imaging predicted recurrence (Fig 2). None of the children with normal vascular imaging had a recurrent stroke, whereas two thirds of those with

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Cerebrovascular Imaging Modalities and Their Findings in Children With AISs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Perinatal ((N = 84))</td>
</tr>
<tr>
<td>Vascular imaging modality</td>
<td></td>
</tr>
<tr>
<td>No vascular imaging</td>
<td>72 (86)</td>
</tr>
<tr>
<td>Vascular imaging(^a)</td>
<td>12 (14)</td>
</tr>
<tr>
<td>Conventional angiography</td>
<td>0 (0)</td>
</tr>
<tr>
<td>MR angiography brain</td>
<td>10 (12)</td>
</tr>
<tr>
<td>MR angiography neck</td>
<td>0 (0)</td>
</tr>
<tr>
<td>MR venography</td>
<td>1 (0)</td>
</tr>
<tr>
<td>CT angiography</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Finding on vascular imaging</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>6/12 (50)</td>
</tr>
<tr>
<td>Large vessel occlusion(^b)</td>
<td>4/12 (33)</td>
</tr>
<tr>
<td>Abnormal vascular imaging(^c)</td>
<td>2/12 (17)</td>
</tr>
<tr>
<td>Large vessel stenosis</td>
<td>2/12 (17)</td>
</tr>
<tr>
<td>Moyamoya syndrome</td>
<td>0/12 (0)</td>
</tr>
<tr>
<td>Vasculitis</td>
<td>0/12 (0)</td>
</tr>
<tr>
<td>Arterial dissection</td>
<td>0/12 (0)</td>
</tr>
</tbody>
</table>

MR indicates magnetic resonance; CT, computed tomography.
\(^a\) Subcategories are not mutually exclusive.
\(^b\) Isolated occlusion, not included in our definition of vascular abnormality.

FIGURE 1

Kaplan-Meier curve of recurrent stroke-free survival among children with perinatal strokes (dashed line, \(n = 84\)) versus later childhood strokes (solid line, \(n = 92\)). This figure demonstrates that children with later childhood strokes are at higher risk of recurrence (\(P = .0003\) by log-rank test).
vascular abnormalities suffered a recurrence within 5 years (Table 3).

Sensitivity Analyses
Receiving vascular imaging was itself associated with an increased risk of recurrence (HR: 5.9; 95% CI: 1.3–26; P = .02), although this association was not significant after excluding 5 children whose initial vascular imaging was performed after their first recurrent event (HR: 3.9; 95% CI: 0.82–18; P = .09). To account for possible selection bias affecting the association between abnormal vascular imaging and recurrence, we excluded these 5 children (3 of whom had additional recurrences after the vascular imaging) from the primary analysis. The remaining children with abnormal vascular imaging had a 5-year cumulative recurrence risk of 57% (95% CI: 33%–84%), again significantly greater than children with normal vascular imaging (P < .0001 by log rank). To assess the possible impact of differential treatment with antithrombotic agents, we compared treatment rates and found them to be similar among children with abnormal (71%) versus normal (77%) vascular imaging (P = .63 by χ² test). Furthermore, treatment with an antithrombotic agent did not predict recurrence (P = .42 by log rank).

DISCUSSION
In an unselected sample from a defined, multiethnic population followed for several years, we found that (1) recurrence rates after childhood AIS are high in an easily identified subgroup, those older than 28 days, (2) the first 6 months are the highest risk period for recurrence, and (3) clinical cerebrovascular imaging is useful for prognostication. Because most children (>28 days old) in our cohort were treated with antithrombotic agents, the true natural history recurrence rates are likely even higher than those that we observed.

Cerebrovascular imaging, although commonly performed in the evaluation of an adult stroke patient, often is not performed in children with AIS. Slightly under half of the cases in our study had such imaging, and there was no change in rates of vascular imaging over the 11 years of study. The only in-print guidelines for the management of later childhood AIS do not explicitly recommend cerebrovascular imaging.9 However, our data demonstrate that these imaging studies can identify those children at highest risk for recurrence: by 5 years, recurrent strokes occurred in none of the children with normal vascular imaging, compared with almost 70% of children with vascular abnormalities. Furthermore, because we based our definition of abnormal vascular imaging on the interpretations of clinical radiologists (many of whom have no formal training in neuroradiology), this finding is generalizable to the typical hospital setting.

Arterial stenosis was the most commonly identified vascular abnormality in our study. It was an isolated finding in 64% of children with stenosis and occurred in the setting of a more diffuse vasculopathy (moyamoya or vasculitis) in the remainder. Although the radiographic appearance of a narrowed vessel lumen can result from a vasculopathy or a recanalizing embolus,

### TABLE 3
Recurrent Stroke Rates Among 52 Children With Later Childhood AIS and Vascular Imaging

<table>
<thead>
<tr>
<th>Finding on vascular imaging</th>
<th>1-y Cumulative Recurrence Rate, % (95% CI)</th>
<th>5-y Cumulative Recurrence Rate, % (95% CI)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal (n = 24)</td>
<td>0 (—)</td>
<td>0 (—)</td>
<td></td>
</tr>
<tr>
<td>Large vessel occlusion (n = 6)b</td>
<td>0 (—)</td>
<td>0 (—)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>Abnormal vascular Imaging (n = 22)c</td>
<td>57 (37–79)</td>
<td>60 (43–87)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Large vessel stenosis (n = 20)</td>
<td>55 (34–78)</td>
<td>64 (41–86)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Moyamoya syndrome (n = 7)</td>
<td>57 (27–90)</td>
<td>78 (41–99)</td>
<td>.003</td>
</tr>
<tr>
<td>Vasculitis (n = 2)</td>
<td>100 (—)</td>
<td>100 (—)</td>
<td>.02</td>
</tr>
<tr>
<td>Arterial dissection (n = 1)</td>
<td>100 (—)</td>
<td>100 (—)</td>
<td>.01</td>
</tr>
</tbody>
</table>

*Log-rank test, compared with normal vascular imaging.

b Isolated occlusion, not included in our definition of vascular abnormality.

c Subcategories of abnormalities on vascular imaging are not mutually exclusive.
the high recurrence risk that we observed with this lesion, similar to primary vasculopathies such as moyamoya, suggests that arterial stenosis behaves more like a true vasculopathy in terms of recurrence. The etiology of isolated large vessel stenosis in children remains unknown and may be heterogeneous. These lesions could be congenital, such as the hypoplastic arteries seen in PHACE (posterior fossa malformations, hemangiomas, arterial anomalies, coarctation of the aorta, cardiac defects, and eye abnormalities) syndrome, or acquired. Many of these cases may represent an entity termed “transient cerebral arteriopathy,” a monophasic focal arteriopathy of childhood involving the large vessels, most often the distal internal carotid artery. However, this diagnosis requires follow-up vascular imaging demonstrating nonprogression, and such imaging was uncommon in our study. Regardless, transient cerebral arteriopathy is a provisional diagnosis that does not specify a pathophysiology, although an infectious etiology (specifically varicella zoster virus) is suspected.

Our definition of abnormal vascular imaging excluded isolated arterial occlusion because such a finding could be present in the absence of a true vasculopathy, such as after a cardioembolic event. It is possible that some children with isolated occlusion had an underlying vasculopathy and were, therefore, misclassified into the “normal” vascular imaging group. Misclassification of a high-risk lesion into the “normal” group could have biased our results toward the null; however, there were no recurrences among the 6 children with isolated arterial occlusion, indicating that no such bias occurred.

Identification of a child at high risk for recurrence on the basis of vascular imaging not only provides prognosis, but also an opportunity for risk reduction. Although there have been no randomized, controlled trials of secondary stroke prevention in children, a variety of strategies are used to reduce the likelihood of a recurrent stroke in the setting of specific high-risk vascular abnormalities. Examples are anticoagulation for arterial dissections, vascular bypass procedures for moyamoya, and immunosuppressive agents for autoimmune vasculitides.

Only 2 previously published studies performed hazard analyses of recurrence after childhood stroke, and both used selected populations in Germany. In a study of perinatal stroke, using a convenience sample and a broader definition of recurrence (including venous sinus thrombosis and deep venous thrombosis), the 5-year cumulative recurrence rate was 3.5%, within the 95% CIs of that found in our study. A study of later childhood stroke, which used a sample derived from hospital populations in various geographic regions of Germany, reported a recurrent stroke rate of 5% within 5 years, which was below the lower limit of our 95% CIs, although their CIs (not reported) may overlap ours. A true difference, however, could be attributable to their lower reported prevalence of vasculopathies (18%).

Our study’s greatest limitation was that vascular imaging was performed only on the subgroup of children whose treating physicians deemed it necessary. Children who did not receive vascular imaging were more likely to have a cardiac or infectious stroke etiology, suggesting that treating physicians may not have pursued vascular imaging once an explanation for the stroke was obtained. Among those who received vascular imaging, the results clearly allowed stratification into high- and low-risk groups. However, repeat vascular imaging was seldom performed, and recent data suggest that progression of cerebrovascular abnormalities may further stratify recurrence risk in children. A additional limitation is that we cannot comment on the preferred type of vascular imaging in children. However, given the high yield of magnetic resonance angiography in children with AIS, as well as the lower risks associated with this noninvasive study, magnetic resonance angiography is probably a reasonable first step.

The strengths of our study include the multiethnic, population-based setting enhancing generalizability, the relatively large sample size given the rarity of this disease, and the extent and quality of follow-up data. In addition, by basing our definition of vascular imaging abnormalities on formal radiology reports, rather than reinterpretation by a study investigator, we were able to demonstrate the prognostic value of such imaging in a typical hospital setting.

We conclude that children with later childhood AIS are at high risk for recurrent stroke, particularly within the first 6 months, and cerebrovascular imaging identifies those children at highest risk. Cerebrovascular imaging should be included in the diagnostic evaluation of children with strokes outside of the neonatal period, a recommendation supported by the Pediatric Stroke Working Group of the Royal College of Physicians in the United Kingdom. It should be emphasized that the high recurrence rates we observed in this study reflect not natural history risk, but rather risk despite current “best medical practice” among a cohort of children with ready access to medical care, many of whom were treated with antithrombotic agents. We need pediatric secondary stroke prevention trials to improve and standardize strategies for reducing recurrent stroke risk in children, particularly among those with abnormal vascular imaging.

ACKNOWLEDGMENTS
This work was supported by an American Heart Association Scientific Development grant and National Institute of Neurological Disorders and Stroke Neurological Sciences Academic Development Award K12 NS01692.

We thank Dr Jean Hayward for thoughtful review of the manuscript.
REFERENCES


17. StataCorp. Stata 9 Reference Manual. College Station, TX: Stata Press; 2005


Auditory Neural Maturation After Exposure to Multiple Courses of Antenatal Betamethasone in Premature Infants as Evaluated by Auditory Brainstem Response

Sanjiv B. Amin, MD, Ronnie Guillet, MD, PhD

Department of Pediatrics, Division of Neonatology, Golisano Children’s Hospital at Strong, University of Rochester, Rochester, New York

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

ABSTRACT

OBJECTIVE. Our goal was to determine if multiple courses of antenatal betamethasone affect auditory neural maturation in 28 to 32 weeks’ gestational age infants.

PATIENTS AND METHODS. A retrospective cohort study was performed to compare auditory neural maturation between premature infants exposed to 1 course of betamethasone and infants exposed to ≥2 courses of betamethasone. Inclusion criteria included all 28 to 32 weeks’ gestational age infants delivered between July 1996 and December 1998 who had auditory brainstem response testing performed (80-dB click stimuli at a repetition rate of 39.9/second) within 24 hours of postnatal life as part of bilirubin-auditory studies. Infants with toxoplasmosis, rubella, cytomegalovirus, herpes infections, chromosomal disorders, unstable conditions, exposure to antenatal dexamethasone, and exposure to ≤1 complete course of betamethasone were excluded. Auditory waveforms were categorized into response types on response replicability and peak identification as types 1 through 4 (type 1 indicating most mature). Absolute and interpeak wave latencies were measured when applicable. Categorical and continuous variables were analyzed by using the χ² test and Student’s t test, respectively.

RESULTS. Of 174 infants studied, 123 received antenatal steroids. Of these, 50 received 1 course and 29 received ≥2 courses of betamethasone. There were no significant differences in perinatal demographics between the 2 groups. After controlling for confounding variables, there was no significant difference in mean absolute wave latencies, mean interpeak latencies, or distribution of response type between the 2 groups. There also was no significant difference in any auditory brainstem response parameters between infants exposed to 1 course of betamethasone (n = 50) and infants exposed to ≥2 courses of betamethasone (n = 17).

CONCLUSION. Compared with a single recommended course of antenatal steroids, multiple courses of antenatal betamethasone are not associated with a deleterious effect on auditory neural maturation in 28 to 32 weeks’ gestational age infants.
Animal studies have demonstrated that repeated courses of antenatal corticosteroids delay myelination and cellular development in the central nervous system. In addition, recent studies using smaller doses of corticosteroids, comparable to those in human clinical trials involving antenatal corticosteroids, have demonstrated alteration of nuclear transcription factors that regulate brain cell differentiation, alteration of neuronal cytoskeleton, and morphologic alteration of synapses. As a result, the second National Institutes of Health consensus development conference in August 2000 on the effect of repeat courses of corticosteroids on fetal maturation emphasized the concern for possible deleterious effects of multiple courses of corticosteroids on the developing central nervous system of premature infants. The consensus report concluded that the use of multiple courses of antenatal corticosteroids should be limited to patients enrolled in clinical research and that clinical studies are required to evaluate the acute and chronic effects of multiple courses of antenatal corticosteroids on the developing brain in premature infants.

Limited existing human data based on observational studies suggest that multiple courses of antenatal corticosteroids are associated with long-term adverse effects on the developing brain. These data also suggest that the long-term adverse effect on the developing brain may be specific to the corticosteroid preparation used, and that repeated courses of betamethasone may not be associated with long-term abnormal neurologic outcome. However, no data exist regarding acute, short-term effects of multiple courses of betamethasone on brain maturation in premature infants.

Auditory brainstem evoked response (ABR) is a non-invasive neurophysiologic assessment of brainstem maturation in premature infants. The ABR waveform in mature neonates is comprised of 3 waves (I, III, and V). Electrophysiologic data have shown that wave I is generated peripherally in the auditory nerve. Wave III reflects the firing of axons exiting the cochlear nuclear complex in the brainstem, whereas wave V primarily reflects an action potential generated by axons from the lateral lemniscus at a more rostral brainstem location. There is a rapid maturation of the ABR that parallels the critical period of brainstem myelination, neuronal development, and axonal growth. With increasing gestational age, maturation of the ABR is characterized by improving detectability of the response peaks and shortening of the absolute wave latencies and interpeak latencies. The decrease in wave I latency reflects the maturation of the peripheral auditory system, whereas the decreases in waves III and V latencies reflect the combined effects of peripheral and central maturation in the auditory system. The decrease in interpeak latencies reflects changes in nerve conduction velocity. Both absolute and interpeak latencies are influenced by degree of myelination, axonal growth, and synaptic function. Because waves I, III, and V are not always detectable in premature infants, the waveform can also be categorized as a response type based on the replicability of the response and the presence of wave III or wave V (Fig 1). The response type also demonstrates progressive maturation with increasing gestational age. There is limited literature about the direct effect of glucocorticosteroids on auditory brainstem responses in humans. In adults, hydrocortisone administration was shown to acutely reduce absolute latencies, but there was no long-term follow-up assessment. We previously demonstrated that there was no significant difference in auditory neural maturation between premature infants who were exposed to antenatal steroids and infants who were not exposed to antenatal steroids as measured by the ABR within 24 hours of birth. However, because there was considerable variation in the total maternal dosage, any dose-dependent effects may have been obscured. This study seeks to determine whether multiple courses of antenatal betamethasone have any effect on auditory...
neural maturation in infants 28 to 32 weeks’ gestational age.

SUBJECTS AND METHODS

Study Design
This was a retrospective cohort study comparing auditory neural maturation of premature infants exposed to multiple courses of antenatal betamethasone with premature infants exposed to a single complete course of betamethasone.

Study Population
All infants 28 to 32 weeks’ gestational age at birth admitted to the NICU of Golisano Children’s Hospital from July 1996 to December 1998 who had an ABR performed within the first 24 hours of postnatal life as part of our ongoing ABR/bilirubin studies were potentially eligible for this study. Infants with craniofacial anomalies, chromosomal disorders, TORCH (toxoplasmosis, other infections, rubella, cytomegalovirus infection, and herpes simplex) infection or those who were too clinically unstable for ABR testing within the first 24 hours after birth were excluded. In addition, infants exposed to antenatal dexamethasone or <1 complete course of antenatal betamethasone therapy (total: 24 mg) were excluded from this study.

Gestational age was assessed by obstetrical dating criteria or, when obstetrical data were inadequate, by Ballard examination. At the time of the study, the recommended antenatal glucocorticoid regimen consisted of the administration to the mothers of two 12-mg doses of betamethasone given intramuscularly 24 hours apart. If the mother had not delivered 1 week after receipt of the betamethasone given intramuscularly, administration to the mothers of two 12-mg doses of betamethasone was repeated. In order to accelerate fetal lung maturation constituted the single course of antenatal betamethasone for any time before delivery, chorioamnionitis, in utero exposure to cocaine and other illicit drugs, use of antenatal magnesium sulfate, 5-minute Apgar score <5, and mechanical ventilation at the time of ABR testing.

ABR
ABRs were recorded with a Biological Navigator evoked-response system, with the subjects lying supine in the isocline and skin temperature >35.5°C. Testing was performed once during the first 12 to 24 hours of postnatal life by audiologists skilled in the administration of ABR tests to NICU infants. Electrode sites were mastoid (reference), midline on high forehead or crown of the head (active), and shoulder (ground). An audiologist inspected the ear canals and removed any visible vernix or debris before each ABR test. Electrode gel was applied to silver/silver chloride electrodes. Bilateral monaural ABR tests were performed by using 80 dB normal hearing level broadband click stimuli with supraaural earphones. The clicks were presented at a repetition rate of 39.9/second, and 3 runs of 2000 repetitions were recorded for each ear. The 2 most replicable runs for each ear were averaged and used for analysis. The ABRs were analyzed by the audiologists without knowledge of gestational age or antenatal steroid status.

Because ABR waves I, III, and V cannot be detected in all premature infants ≤32 weeks’ gestational age, ABR waveforms were categorized into response types on the basis of response replicability and peak identification: type 1, a waveform with normal morphology and replicable waves III and V; type 2, a replicable response with

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Demographic Profile of the Study Population as a Function of 1 or ≥2 Courses of Antenatal Betamethasone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Course of Betamethasone</td>
<td>≥2 Courses of Betamethasone</td>
</tr>
<tr>
<td>(N = 50), Group I</td>
<td>(N = 29), Group II</td>
</tr>
<tr>
<td>Birth weight, g*</td>
<td>1396 ± 272</td>
</tr>
<tr>
<td>Gestational age, wk**</td>
<td>30.2 ± 1.2</td>
</tr>
<tr>
<td>Small for gestational age, %</td>
<td>6</td>
</tr>
<tr>
<td>Gender, % male***</td>
<td>46</td>
</tr>
<tr>
<td>Race, % black****</td>
<td>34</td>
</tr>
<tr>
<td>Exposure to magnesium sulfate, %</td>
<td>70</td>
</tr>
<tr>
<td>Maternal chorioamnionitis, %</td>
<td>24</td>
</tr>
<tr>
<td>In utero exposure to illicit drugs, %</td>
<td>10</td>
</tr>
<tr>
<td>Rate of delivery by cesarean section, %</td>
<td>44</td>
</tr>
<tr>
<td>Apgar score &lt;5 at 5 min, %</td>
<td>2</td>
</tr>
<tr>
<td>Mechanical ventilation on day 1, %</td>
<td>56</td>
</tr>
</tbody>
</table>

* Mean ± SD using t test.
** Proportions were analyzed by using the χ² test.
*** The P value was significant when compared with 1 course of betamethasone.
either a wave III or V; type 3, a replicable response with neither a wave III or V; and type 4, a waveform with no replicable response (Fig 1). If the waveform was type 1 or 2, latencies for waves I, III, and V and interpeak latencies I to III, III to V, and I to V were measured. The response for the better ear was used for final analysis. The study was approved by the Human Subject Review Board, and informed consent was obtained from parents before ABR testing.

Sample-Size Calculations and Statistical Analysis
An approximate sample size was determined for the number of neonates to be studied on the basis of earlier findings of ABR maturation study. On the basis of earlier findings, 15 subjects in each group would allow detection of actual difference of 0.7 milliseconds (equal to \( \frac{\text{0.75 SD}}{\sqrt{0.7}} \)) for absolute latencies and interpeak latencies for infants 28 to 32 weeks’ gestational age with an \( \alpha \) level of .05 and a power of .80.

Student’s \( t \) test was used to analyze continuous variables using Stata (Stata Corp, College Station, TX). A \( \chi^2 \) or Fisher’s exact test, as appropriate, was used to analyze nominal variables. Multiple regression analysis was performed to assess the independent relations of multiple variables with measured ABR parameters. All tests were 2 sided, and a \( P \) value of <.05 was considered statistically significant.

RESULTS
Demographics
A total of 174 infants 28 to 32 weeks’ gestational age were eligible for this study. Of these, 123 received antenatal steroids, whereas 51 infants did not. Of 123 infants who received antenatal steroids, 29 infants who received <1 complete course of betamethasone, 5 infants who received antenatal dexamethasone therapy, and 6 infants who received between 1 and 2 courses of betamethasone were excluded. For 4 outborn infants, it was unclear whether infants received a partial or complete course of antenatal steroid therapy, thus they were also excluded. Of the remaining 79 infants, 50 infants (group I) received 1 complete course of antenatal betamethasone, whereas 29 infants (group II) received \( \geq 2 \) courses of antenatal betamethasone. Of 29 infants, 17 infants (group IIb) received \( \geq 2 \) courses of betamethasone (mean dosage: 80 mg; range: 60–144 mg). Among group I infants, 95% of infants received the last dose of betamethasone \( > 24 \) hours but less than a week before delivery. Among group II infants, 100% of infants received the last dose of betamethasone \( > 24 \) hours but less than a week before delivery. The demographics of the study patients as a function of exposure to 1 or \( \geq 2 \) courses of antenatal betamethasone are shown in Table 1. There were no significant differences between group I and group II in gestational age at birth (\( P = .4 \)), birth weight

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Absolute Latencies and Interpeak Latencies as a Function of 1 or ( \geq 2 ) Courses of Antenatal Betamethasone in Infants 28 to 32 Weeks’ Gestational Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Antenatal Steroids</td>
<td>Group I, Mean ± SD</td>
</tr>
<tr>
<td>( &gt; 2 ) Courses of Betamethasone</td>
<td>( &gt; 2 ) Courses of Betamethasone</td>
</tr>
<tr>
<td>(N = 50); Group I</td>
<td>(N = 29); Group II</td>
</tr>
<tr>
<td>Absolute latencies, msec</td>
<td>Latency I</td>
</tr>
<tr>
<td></td>
<td>Latency III</td>
</tr>
<tr>
<td></td>
<td>Latency V</td>
</tr>
<tr>
<td>Interpeak latencies, msec</td>
<td>Latency I-III</td>
</tr>
<tr>
<td></td>
<td>Latency III-V</td>
</tr>
<tr>
<td></td>
<td>Latency I-V</td>
</tr>
</tbody>
</table>

NS indicates nonsignificant.
small for gestational age (P = .6), gender distribution (P = .6), race (P = .1), exposure to antenatal magnesium sulfate (P = .7), maternal chorioamnionitis (P = .9), in utero exposure to illicit drugs (P = .3), rate of cesarean-section delivery (P = .4), 5-minute Apgar score <5 minutes (P = .4), or use of mechanical ventilation (P = .2) at the time of testing.

Wave Latencies
There were no significant differences in absolute latencies and interpeak latencies between group I and group II (P > .23; Table 2). The differences in absolute latencies and interpeak latencies between the infants in the 2 groups remained insignificant when controlled for gestational age, birth weight, race, maternal treatment with magnesium sulfate, chorioamnionitis, in utero exposure to illicit drugs, and 5-minute Apgar score <5 using a multiple regression model. For comparison, absolute latencies and interpeak latencies for infants 28 to 32 weeks’ gestational age (mean ± SD: 30.6 ± 1.2 weeks) who did not receive antenatal steroids (n = 51; Table 2) are given. There was no significant difference in absolute latencies and interpeak latencies between infants who did not receive antenatal steroids and infants who received 1 or multiple courses of antenatal betamethasone.

Response Types
The frequency distribution of response types are shown in Table 3. The difference in the frequency distribution of response types between the infants in group I and group II remained insignificant (P = .7) after controlling for possible confounders, as above.

Subgroup Analysis
On subgroup analysis of the data for infants who received >2 courses of betamethasone (group IIb) compared with infants who received 1 course of betamethasone (group I), there was no difference in any of the demographic characteristics except for race and use of mechanical ventilation on the first day (Table 1). Infants exposed to >2 courses of antenatal betamethasone were less likely to be mechanically ventilated on day 1 (P = .05). No differences in absolute or interpeak latencies between these subgroups (P > .25; Table 2) were found even after controlling for potential confounding factors.

There was also no difference in the frequency distribution of response types (P = .5; Table 3).

In additional analysis to evaluate the effect of time interval from the last dose of antenatal steroid, there was no significant difference (P > .23) in absolute latencies and interpeak latencies between infants who were exposed to the last dose of betamethasone >72 hours before delivery (n = 19) and infants who were exposed to the last dose of betamethasone ≤72 hours before delivery (n = 60).

There was also no linear correlation between absolute latencies or interpeak latencies and amount of maternal betamethasone dosage (R² ≤ 0.09; P > .22)

DISCUSSION
Both animal studies and observational studies of premature infants suggest that repeated courses of antenatal steroids may result in long-term neurodevelopmental impairment.3,4 As a result, the second National Institutes of Health consensus development conference in August 2000 concluded that the use of multiple courses of antenatal corticosteroids should be limited to patients enrolled in clinical research studies. The only published randomized, controlled trial of such treatment was stopped early because of the concern raised by the consensus group.4 Approximately half of the original number of patients was enrolled before closing the study. The investigators found that the infants of women who received multiple antenatal courses of steroids had less severe respiratory distress syndrome (RDS), but there was a trend toward an increase in the incidence of intraventricular hemorrhage and chorioamnionitis. Because of the apparent lack of efficacy for the primary outcome and the decision to “do no harm,” the study was terminated early. No other trials have been completed to date.

The use of the ABR as a surrogate marker for central nervous system effects of multiple courses of antenatal steroids is based on the demonstrated effects of steroids on neuronal differentiation and synapse formation.2,3 The ABR is a noninvasive assessment of brainstem maturation in premature infants. The 3 waves that comprise the waveform represent activity at different levels of the auditory pathway. By evaluating the wave latencies, which are influenced by the degree of myelination, axonal growth, and synaptic function, inference can be

<p>| TABLE 3 Frequency Distribution of ABR Response Types as a Function of Exposure to 1 or ≥2 Courses of Antenatal Betamethasone |</p>
<table>
<thead>
<tr>
<th>Response Types</th>
<th>No Antenatal Steroids (N = 51), %</th>
<th>1 Complete Course of Betamethasone (N = 50), Group I, %</th>
<th>≥2 Courses of Betamethasone (N = 29), Group II, %</th>
<th>&gt;2 Courses of Betamethasone (N = 17), Group IIb, %</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>42</td>
<td>48</td>
<td>45</td>
<td>65</td>
<td>NS</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>16</td>
<td>10</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>22</td>
<td>24</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>14</td>
<td>21</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

NS indicates nonsignificant.
made about the possible effects of antenatal steroids. As in the published randomized, controlled trial, we could find no evidence for either a beneficial or a harmful effect of multiple courses of betamethasone in utero brainstem maturation compared with single courses.

By using ABR, we previously demonstrated that antenatal steroid therapy is not associated with acute adverse effects on auditory neural maturation in infants 24 to 32 weeks’ gestational age. However, there was considerable variation in the total maternal dosage that might have obscured any dose-dependent effects. In our analysis, a trend toward an increased prevalence of more mature response types was found in infants who received >2 courses of antenatal betamethasone compared with those who received fewer courses. There was no effect, either beneficial or deleterious, on absolute and interpeak latencies. This may reflect a similar lack of effect on maturation of other parts of the brain.

These findings are inconsistent with the findings from animal studies. One explanation may be the difference between humans and laboratory animals in the timing of the brain growth spurt and the complexity of brain development. In humans, most of the neuronal division is completed by 24 weeks’ gestation. After 24 weeks, cell division in the brain involves mainly the oligodendroglial cells that will lay down myelin. Secondly, the adverse effects on developing brain may be specific to the corticosteroid preparation and may not be associated with betamethasone. French et al in an observational study involving premature infants showed that multiple courses of betamethasone were not associated with abnormal neurologic outcome in premature infants at 3 years of age. Similarly, Baud et al in a retrospective cohort study reported that ≥1 courses of antenatal betamethasone was associated with a lower risk of cystic periventricular leukomalacia compared with no antenatal steroid therapy or ≥1 courses of antenatal dexamethasone. Recently, Spinillo et al in a prospective observational study reported that the risk of periventricular leukomalacia and abnormal neurodevelopmental outcome in premature infants was associated with multiple courses of dexamethasone but not to multiple courses of betamethasone.

Antenatal steroids are used primarily for their beneficial effect on lung maturation. Although the only published randomized trial suggests that multiple courses of antenatal steroids in infants 24 to 32 weeks’ gestational age are not beneficial overall, subgroup analysis demonstrated that multiple courses were associated with beneficial pulmonary effects in infants ≥28 weeks’ gestation. The meta-analysis of observational studies in premature infants also found that multiple courses of antenatal steroids were associated with a decreased incidence of RDS and patent ductus arteriosus. Similarly, we found that fewer infants who were exposed to multiple courses required mechanical ventilation on the first day compared with infants exposed to 1 course of betamethasone.

The major limitation of our study is its retrospective nature. However, ABR measurements and response type assignments were performed by audiologists without knowledge of the infant’s antenatal steroid status. Because our analyses were limited to infants who were exposed in utero to betamethasone, the findings may not be generalizable to premature infants exposed to dexamethasone. Although congenital middle ear effusion may affect absolute latencies, the prevalence of congenital middle ear effusion as evaluated by tympanometry or pneumatic otoscope is low (<8%) in term neonates. In premature infants, it is technically difficult to accurately evaluate middle ear disease by using tympanometry or pneumatic otoscope. We, therefore, used 80 db to decrease the interference from middle ear disease and external noise sources, including noise from any form of respiratory support. All subjects passed otoacoustic emission test at 34 to 35 weeks’ postmenstrual age, indicating normal middle ear function.

Findings from this study suggest that ABR may be a surrogate outcome marker that can be used to assess the potential effect of prenatal exposure to steroids or other drugs with central nervous system interactions. A decreased incidence of RDS and patent ductus arteriosus, in the absence of either short- or longer-term adverse neurologic effects, suggests that randomized trials of multiple courses of betamethasone in premature infants should be considered.

ACKNOWLEDGMENTS
This research was partially supported by National Institutes of Health grant K-23 DC 006229-02.

We thank Mark Orlando, PhD, Kathleen Merle, MS, Todd M. Gibson, AuD, Teri D. Holt, MS, Matthew MacDonald, AuD, Lynette McRae, MA, Diane S. Puccia, MA, and Catherine Papso Seeger, MS, for performing ABR on premature infants. We also thank Kristina Mossgraber for collecting maternal information.

REFERENCES
proteins MAP1B and MAP2 in fetal sheep. *J Physiol.* 2001;530:497–506


---

**CHINA: 2 MILLION TEENAGERS ADDICTED TO INTERNET**

“Chinese teenagers are becoming addicted to the Internet, the state news media reported. Of China’s 18.3 million teenage Internet users, more than 2 million are already addicts, The China Daily said, citing a study by the Communist Youth League. Quoting Gao Wenbin, a psychology researcher at the Chinese Academy of Sciences, the article went on: ‘Internet addicts in China are as many as 10 years younger than those in the West. They are more susceptible.’ Most are males between 15 and 20, he said, with as many as 15 percent needing ‘urgent help.’ He blamed a lack of diversions. ‘They will naturally turn to the virtual world if they cannot find an outlet for their energy,’ he said.”


Noted by JFL, MD
Stress Predicts Brain Changes in Children: A Pilot Longitudinal Study on Youth Stress, Posttraumatic Stress Disorder, and the Hippocampus

Victor G. Carrion, MD*, Carl F. Weems, PhD*, Allan L. Reiss, MD*

*Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California; Department of Psychology, University of New Orleans, New Orleans, Louisiana

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

ABSTRACT

OBJECTIVE. Does stress damage the brain? Studies of adults with posttraumatic stress disorder have demonstrated smaller hippocampal volumes when compared with the volumes of adults with no posttraumatic stress disorder. Studies of children with posttraumatic stress disorder have not replicated the smaller hippocampal findings in adults, which suggests that smaller hippocampal volume may be caused by neurodevelopmental experiences with stress. Animal research has demonstrated that the glucocorticoids secreted during stress can be neurotoxic to the hippocampus, but this has not been empirically demonstrated in human samples. We hypothesized that cortisol volumes would predict hippocampal volume reduction in patients with posttraumatic symptoms.

PATIENTS AND METHODS. We report data from a pilot longitudinal study of children (n = 15) with history of maltreatment who underwent clinical evaluation for posttraumatic stress disorder, cortisol, and neuroimaging.

RESULTS. Posttraumatic stress disorder symptoms and cortisol at baseline predicted hippocampal reduction over an ensuing 12- to 18-month interval.

CONCLUSIONS. Results from this pilot study suggest that stress is associated with hippocampal reduction in children with posttraumatic stress disorder symptoms and provide preliminary human evidence that stress may indeed damage the hippocampus. Additional studies seem to be warranted.
For more than a decade, a medial temporal brain structure named the hippocampus, because of its sea-horse shape, has been studied in the pathophysiology of posttraumatic stress disorder (PTSD). The hippocampus’ functional role in memory processing and its anatomic location (part of the limbic system) has made this structure a prime candidate for investigation in relation to PTSD because the disorder is characterized by intrusive recollections of the traumatic event and difficulties with emotion regulation. Research into the neurobiology of PTSD has been aided by the availability of high-resolution MRI, which has made it possible to obtain accurate volumetric measurements of the hippocampus. The literature reports significantly reduced hippocampal volumes in adults with PTSD when compared with healthy controls.1–5 These findings have not been replicated when investigating the hippocampal volume of children with PTSD.6–8 These studies have found either no difference between PTSD and controls in hippocampal volume6,7 or larger hippocampal volumes secondary to larger volumes of white matter, but no gray matter.8

The relationship between the hippocampus and PTSD remains to be determined because many questions are unresolved. Why have findings failed to replicate in youth? Are the identified volumetric abnormalities of the hippocampus in this condition a risk factor for developing symptoms of PTSD or are the structural differences a biological marker of stress for this behaviorally defined disorder? Pathophysiological differences between adult and pediatric PTSD may, in part, be explained by clarifying the role of stress in human hippocampal development.

The hippocampus has a protracted ontogeny, persistent postnatal neurogenesis, and a high density of glucocorticoid receptors.9–11 An increasing body of research suggests that stress and the glucocorticoids secreted by the adrenal steroids during stress can damage the hippocampus.12,13 This damage includes alteration of pyramidal cell morphology, pyramidal cell death, and suppression of granule cells.14,15 These stress-mediated hippocampal changes can also be induced by injection of corticotropin-releasing factor.15 Because of our work16 and others17 documenting high levels of cortisol in children with PTSD, we reasoned that it would be important to study the associations among cortisol levels, PTSD symptoms, and hippocampal volume changes in children with a history of traumatic stress. We have found that prebedtime salivary cortisol levels demonstrated the largest difference in values between children with PTSD and controls (ie, it seems to be the best marker of pathogenic cortisol levels).

The risk-versus-marker question remains unanswered because to date, most of the studies examining hippocampal volume have been cross-sectional in design. De Bellis and colleagues18 assessed hippocampal volumes in 9 prepubertal maltreated children and 9 healthy controls at baseline and after 2 years’ follow-up. They found no differences in hippocampal changes between groups at baseline, follow-up, or across time. They noted, however, that left hippocampal volume decreased across time in both groups, and that this may have been influenced by the sample’s postpubertal status. Group comparisons may have been underpowered and may not have adequately explored hippocampal change within the PTSD group. Such groups of youth are often quite heterogeneous.

Given the above findings, we theorized that hippocampal volume reductions in adults may result in part from chronic exposure to stress throughout childhood development. A lifetime accumulation of stress effects on the hippocampus would help explain the positive volumetric findings in adults and the lack of findings in younger samples. The impact of chronicity was studied by Wignall and colleagues19 who found smaller right hippocampal volumes in adults with recent-onset PTSD after acute trauma, suggesting that either hippocampal damage occurs within months of the trauma or that smaller hippocampal volume is a predisposing factor for PTSD. When examined longitudinally, however, individuals with acute trauma assessed at baseline and who developed PTSD 6 months later (n = 10) did not differ from those individuals who did not develop PTSD (n = 27) in terms of hippocampal size at either baseline or at follow-up.20 This suggests that smaller hippocampal volume is not a risk factor for developing PTSD after an acute event. In a pivotal monozygotic twin study, Gilbertson and colleagues21 compared twins who were discordant for chronic trauma (Vietnam combat) exposure. They found that unexposed co-twins of those with PTSD also had smaller hippocampi, suggesting that smaller hippocampal volume is a preexisting familial vulnerability factor. Although this study does not rule out the possibility of an added toxic environmental effect, they note that those siblings with combat-exposure did not have any additional reduction of their hippocampal volumes.

Differences between findings in adult and pediatric studies regarding hippocampal volume may also be attributed to neuromaturational factors. Early life stress may alter synaptogenesis, dendritic proliferation, and pruning, but these effects may not manifest until later development. For example, preclinical studies have shown that differences in hippocampal synaptic density arise in rats exposed to early stress only postpubertally as a consequence of attenuated synaptogenesis.22

On the basis of the theory that the changes in hippocampal volume found in the adult PTSD brain result from prolonged exposure to neurotoxic levels of cortisol secreted during chronic stress, we conducted a longitudinal study of children with history of traumatic stress who already show signs and symptoms of PTSD.
evaluation entails behavioral, endocrinological, and neuroanatomical (via structural MRI) assessment at baseline (time 1) and follow-up (time 2; 12–18 months later). Drawing from the existing literature and our hypothesis that hippocampal volume reductions may result, in part, from chronic exposure to stress throughout childhood development, we hypothesized that PTSD symptoms and cortisol levels would be associated with changes in hippocampal volumes across time 1 to time 2. We focused on prebedtime cortisol levels (for the reason noted above) and total PTSD symptoms and explored the specific role of hyperarousal symptoms (cluster D). Of the 3 clusters of PTSD [reexperience (B), avoidance and numbing (C), and hyperarousal (D)], we have found D to be highly pathognomonic of PTSD in childhood and a strong predictor of later impairments.23–25

PATIENTS AND METHODS

Participants
Participants were recruited from local departments of social services and mental health clinics, and all of the participants fulfilled the following criteria: (1) at least 1 episode of exposure to trauma, as defined by Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criterion A1: “the person experienced, witnessed, or was confronted with an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others”;26 (2) age 7 to 13 years old; (3) a severity score of ≥12 on the PTSD Reaction Index. Exclusion criteria consisted of a history of mental retardation; history of schizophrenia or autism, presence of any metal or electrical conductive implants or foreign bodies; and current history of subarachnoid hemorrhage, head trauma, epilepsy, or other documented neurologic disorder. Most children experienced multiple traumatic events. Traumatic events included witnessing violence, physical abuse, separation and loss, sexual abuse, physical neglect, and emotional abuse.

The sample consisted of 6 boys and 9 girls, for a total sample of 15 children. The mean age of the children was 10.4 years, with a range of 8 to 14 years. In terms of family income, 47% reported incomes between 0 and $21,000, 27% reported incomes between $21,000 and $41,000, and 13% of the families reported incomes over $41,000 (2 families did not report income). Ethnic composition was white (n = 7), black (n = 6), Hispanic (n = 1), and Asian (n = 1). Children’s average pubic hair Tanner stage was 2.1; for girls average breast Tanner stage was 2.0, for boys average genital Tanner stage was 2.1. Wechsler Abbreviated Scale of Intelligence full-scale IQs ranged from 62 to 142, with an average score of 89.0. (Two participants scored below 70. These participants were included in the sample because they had no previous history of mental retardation and were able to participate in the clinical evaluation).

Clinical Evaluation
All participants and their legal guardians were presented with an institutional review board-approved informed consent and agreed to participate. An in-depth clinical evaluation was conducted on all referred children with an PTSD Reaction Index score ≥12.

Evaluation Instruments
To assess PTSD symptoms, we used the Clinician-Administered PTSD Scale for Children and Adolescents (CAPS-CA).27 The CAPS-CA is a developmentally sensitive counterpart to the CAPS for adults.28 This clinical interview consists of standardized prompt questions, supplementary follow-up (probe) questions, and behaviorally anchored 5-point rating scales corresponding with the frequency and intensity of each symptom assessed. The CAPS-CA assesses all DSM-IV criteria for PTSD. Specifically, it assesses exposure to criterion A1 events and the individuals’ experience of these events (ie, criterion A2), the 17 symptoms for PTSD clustered in DSM-IV (ie, criteria B, C, and D), and the 1-month duration requirement (criterion E). The frequency and intensity of each of the symptoms were rated via behaviorally anchored ratings (0–4 for frequency and 0–4 for intensity) and then summed. The CAPS-CA has good internal consistency estimates for the ratings and has shown concurrent validity with the Child PTSD Checklist.29 A certified child psychiatrist (Dr Carrion) who was trained on the administration of instrument conducted all CAPS-CA interviews. The following variables were computed with the CAPS-CA: (1) total traumatic stress symptoms and (2) hyperarousal symptoms: a composite score of ratings from the 5 hyperarousal items from symptom cluster D.

Biological Maturation
Participants’ pubertal development was determined by self-report. Participants selected from drawings with written descriptions representing the 5 Tanner stages of pubic hair development and genital development for boys and breast development for girls. Previous research has demonstrated that self-report Tanner staging is a valid and reliable method that has been shown to correlate with physician ratings.30

Neuroendocrine Evaluation
Salivary Cortisol
Cortisol samples were obtained from the participant during home-based measurements. They were collected 4 times a day (prebreakfast, prelunch, predinner, and prebedtime) over the course of 3 consecutive days producing 12 samples. Detailed instructions were provided
to caretakers and children regarding the collection of saliva samples. Samples were collected by placing a cotton swab in the participant’s mouth for 1 minute. The cotton was then placed inside a sterile plastic tube, sealed, and kept refrigerated. Saliva was extracted from the cotton by centrifuging the plastic tubes and cotton for 8 to 10 minutes. The cotton was then removed and the tubes sealed. All samples were kept at −20°C and shipped on dry ice to the laboratory for assay. Samples were processed using the Magic Cortisol radioimmunoassay kit produced by Ciba-Corning (Giessen, Germany) as adapted for salivary cortisol analysis by the University of Minnesota Endocrine Laboratory.32 Interassay and intraassay coefficients of variation are maintained at <12%. Cortisol is reported in micrograms per deciliter. As recommended for increased reliability,33 an aggregate score (mean) from the 3 days was created for each time period, and to reduce the number of statistical tests we used the mean cortisol level for the aggregated cortisol levels for analyses in this investigation. Individuals showed relative stability in their cortisol levels across the 3 days. Specifically, the correlation between day 1 and day 2 mean cortisol level was \( r = 0.46; P < .001 \) between day 1 and day 3 \( r = 0.52; P < .001 \) and day 2 and day 3 \( r = 0.31; P < .001 \). Prebedtime cortisol was the primary variable used on the basis of our previous findings identifying this value as the largest difference between children with PTSD and healthy controls.16

**Neuroanatomical Evaluation**

**MRI Acquisition**

MRI data were acquired using a 1.5-T GE-Signa scanner (General Electric, Milwaukee, WI). Coronal 3-D volumetric spoiled gradient echo series (repetition time: 35 milliseconds; echo time: 6 milliseconds; flip angle: 45°; number of excitations: 1; field of view: 24; matrix: 256 × 192), and 124- to 1.5-mm contiguous slices were acquired on all participants and were used for all measurements and analysis.

**Image Analysis**

Morphometric analysis was performed at the Center for Interdisciplinary Brain Sciences Research at Stanford University and was conducted by research staff blind to the hypothesis in this study and the PTSD status of the participant. Volumetric assessment of segmented image data in the software program, BrainImage,34 requires a stepwise process of data importation, removal of non-brain voxels, correction of image nonuniformity, positional normalization, and fuzzy tissue segmentation. The assessment of hippocampus volumes in BrainImage required a manual delineation of regions of interest. Brain tissue was isolated, and coronal images were oriented perpendicular to the anterior commissure-posterior commissure plane. Interrater reliability testing was performed between 2 well-trained raters blind to diagnoses to ensure accuracy in measurements. A single rater then circumscribed regions of the hippocampus for all participants on coronal images oriented perpendicular to the anterior commissure-posterior commissure plane and according to a protocol previously developed in our laboratory.35 To increase the resolution at which the regions of interest could be drawn, the matrix sizes of the coronal data sets were expanded from 256 to 512 × 10^2 pixels by using a bicubic interpolation algorithm. The image contrast was increased so that the amygdala and hippocampus were clearly distinguishable from the surrounding white matter and cerebrospinal fluid (CSF). Volume measures recorded total tissue and were examined separately for right and left volumes. The anteriormost slice of the hippocampus was determined by the presence of the alveus and by the development of a laminar structure that distinguishes the hippocampus from the amygdala. The borders were defined by the surrounding white matter and CSF and superiorly by the amygdala. The hippocampus excluded the tail of the caudate nucleus anteriorly and excluded the thalamus posteriorly. Circumscription of the hippocampus continued until it disappeared posteriorly, approximately at the point where the corpus callosum fuses with the fornix. Time 1 and time 2 hippocampal volumes were highly correlated (left \( r = 0.89; \) right \( r = 0.85 \)).

**Statistical Methods**

The association between changes in hippocampal volumes and PTSD symptoms was examined by using Pearson correlations, and between hippocampal volumes and cortisol were examined by using Spearman correlations because of the nonnormal distribution of prebedtime cortisol levels using unidirectional tests at \( P < .05 \). The association was examined using a 2-stage process. The change in volume was first calculated by simple subtraction of time 1 volumes from time 2 volumes. Thus, positive numbers represented increases in volume and negative numbers decreases in volume. In addition to testing simple change values, we also conducted a more stringent test of the associations among cortisol, PTSD symptoms, and hippocampal development. Specifically, we examined change in hippocampus volumes from time 1 to time 2 while controlling for biological maturation (Tanner stage) and gender. To do this, a change variable was computed by calculating the standardized residuals of time 2 hippocampus volumes predicted by Tanner stage, gender, and time 1 hippocampus volumes via regression analyses. These standardized change scores thus represent the difference in expected time 2 hippocampal volume on the basis of Tanner stage, gender, and time 1 hippocampus volumes. Positive scores represent a relative increase, whereas negative scores represent a relative decrease. This technique and the unidirectional tests, as opposed to a multiple linear
RESULTS

Examination of the score ranges and skew indicated that cortisol levels were not normally distributed and that there was 1 significant outlier in change in right hippocampus volumes (the score was a full SD increase from the next closest increase). This case was excluded from the analyses. Descriptive statistics for the variables of interest are presented on Table 1.

As noted, hippocampal change and residualized change were computed such that a negative number indicates a decrease in predicted volume over time. Correlations among the measures of PTSD symptoms, cortisol, and hippocampal change are presented in Table 2 and indicated that the severity of PTSD symptoms and the cortisol levels at time 1 were significantly negatively correlated with change in right hippocampus volumes using both the simple change score, as well as the residualized change score. No significant correlations were found for change in left hippocampus volumes. The association between residual change in right hippocampus volumes and PTSD symptoms and cortisol levels are depicted in Fig 1.

Exploratory analyses were conducted to determine whether hippocampal volumes at time 1 were associated with change (calculated similarly as hippocampus change) in PTSD symptoms from time 1 to time 2, and results indicated that this association was not statistically significant ($r = -0.04$; $P = .45$).

DISCUSSION

Our results support the hypotheses that PTSD symptoms and cortisol levels at baseline are associated with changes in hippocampal volume over an ensuing 12- to 18-month interval. Specifically, we found that severity of PTSD symptoms and cortisol levels predict a reduction in hippocampal volume from baseline to follow-up when controlling for pubertal maturation and gender in children with a history of traumatic stress. This is the first longitudinal study in PTSD to document an association between hippocampal changes with PTSD symptoms and with a marker of stress, cortisol levels. These longitudinal findings help elucidate previous cross-sectional reports of smaller hippocampal volumes in PTSD populations. Our results are also in accord with animal literature reporting on the neurotoxic effects of glucocorticoids in the hippocampus. Our results stand in contrast, however, with studies identifying hippocampal volume as a vulnerability factor. Although, this study was not designed to address the vulnerability factor hypothesis, our exploratory analyses suggest that hippocampal volume was not a risk factor for development of PTSD symptoms.

Participants with the highest severity of PTSD symptoms, and more specifically hyperarousal symptoms, had a reduction of their hippocampal volume. Behavioral manifestations of the syndrome may have direct developmental effects on a brain structure involved in managing cognitive processes. In past research, we have found that hyperarousal symptoms in children with PTSD predict the development of cognitive difficulties in this syndrome. Such findings may help elucidate the reasons for cognitive impairment in PTSD, and this mechanism speaks to the inherent chronicity of the disorder because less cognitive resources could jeopardize recovery from PTSD.

regression strategy for example, were chosen to balance type 1 and 2 error rates because of the relatively small sample size. Specifically, using a single regression analysis (as opposed to our 2-stage process) with 4 predictors only leaves 9 degrees of freedom, making it virtually impossible to detect a significant effect with the present sample size (in other words, it inflates type 2 error beyond reasonable expectations). With a sample size of 14 or 15, this sets as “significant” findings where the association explains 20% or more of the covariance, a fairly large effect size for these type of studies, thus balances type 1 (ie, conclude there is an effect when there is none) and type 2 error (ie, conclude there is not an effect when there is one).
Theoretically, PTSD symptoms can enhance stress in a variety of ways. For example, by interfering with the normal sleep cycle and with emotion regulation, hyperarousal symptoms may overactivate the stress-response system. This increase in noradrenergic and adrenaline response, as well as in activity of the hypothalamic-pituitary-adrenal (HPA) axis, which results in cortisol secretion, may strain the system into what has been referred to as an allostatic load. This allostatic load can have a direct effect on the function of these stress-response systems by altering their specificity to respond on exposure to additional stress.
Our cortisol findings address a potential mechanism by which stress can alter the hippocampus. There is substantial animal literature demonstrating the neurotoxic effects of glucocorticoid in the glucocorticoid receptor-rich hippocampus.\textsuperscript{14,38,39} Glucocorticoids can also exert their neurotoxicity indirectly via accumulation of extracellular glutamate.\textsuperscript{40} High levels of glucocorticoids have been reported in children with history of maltreatment and PTSD.\textsuperscript{16,17} Elevated cortisol levels suggest that high levels of stress lead to activation of the HPA axis and cortisol production and that this leads to hippocampal toxicity, which results in poor inhibitory activity from the hippocampus unto other centers, such as the HPA axis itself. The putative neurotoxic effects of cortisol on the hippocampus may depend on at least 3 factors: (1) the developmental stage of the structure (the hippocampus glucocorticoid receptors density may change throughout development), (2) the level and sustainability of cortisol released, and (3) the severity and/or chronicity of the stressful events.

Most studies in adults that have found a reduced hippocampal volume in PTSD have studied populations with chronic symptoms or chronic trauma (Wiglall et al\textsuperscript{19} is an exception). A longitudinal study on acute trauma and PTSD found no hippocampal differences.\textsuperscript{20} Although our participants were children, their histories of severe, chronic maltreatment may have already caused enough stress early in life, so that these hippocampal changes can be identified. On the other hand, Gilbertson and colleagues\textsuperscript{21} who were also studying individuals with chronic history (Vietnam combat) did not find support for the theory of environmental toxicity from stress. The time since trauma, however, may also impact the volume of the hippocampus. Although the participants continue to have symptoms of PTSD, there may exist enough time for the plasticity of the hippocampus to undergo at least partial recovery. Such a conclusion is consistent with our findings should be considered.

In addition to the possibilities that a small hippocampus is a marker of PTSD or that a small hippocampus is a vulnerability factor for PTSD, other alternatives that are consistent with our findings should be considered. For example, independent of PTSD status, a small hippocampus may be a marker of chronic or ongoing stress. In addition, there can exist genetic and/or familial risk factors for enhanced vulnerability to the neurotoxic agents from stress. Finally, there can be a synergistic effect between a small hippocampus and the environmental neurotoxicity of stress. Our data could support all of these possibilities, with the exception of a small hippocampus being solely a risk factor for PTSD.

The answer to the role of the hippocampus in PTSD is complex. This study reports on an association between stress, cortisol, PTSD, and hippocampal change that starts early in life. Although this sample represents the largest longitudinal study of hippocampal volumes in traumatized youth to date, the study and range of conclusions about additional confounding variables is limited by the sample size and findings should be considered preliminary until replicated. However, future structural imaging studies should conduct longitudinal investigations in individuals at risk for trauma before and after experiencing trauma and PTSD symptoms. It will be important to be able to accurately measure trauma type and duration, as well as duration of PTSD symptoms.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health grant MH63893-01, the National Alliance for Research on Schizophrenia and Depression, the American Foundation for Suicide Prevention, Aloha Foundation grants (to Dr Carrión), and National Institutes of Health grants MH01142, MH19908, MH050047, and HD31715 (to Dr Reiss).

REFERENCES

12. Sapolsky RM. Potential behavioral modification of glucocorti
39. Sapolsky RM. Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Arch Gen Psychiatry.* 2000;57:925–935
Twelve-Month Effectiveness of a Parent-led, Family-Focused Weight-Management Program for Prepubertal Children: A Randomized, Controlled Trial

Rebecca K. Golley, BSc, BNutrDiet, PhD; Anthea M. Magarey, BSc, DipNutrDiet, PhD; Louise A. Baur, BSc, MBBS, PhD; Katharine S. Steinbeck, MBBS, PhD; Lynne A. Daniels, BSc, DipNutrDiet, MSc, PhD

ABSTRACT

BACKGROUND. Parenting-skills training may be an effective age-appropriate child behavior-modification strategy to assist parents in addressing childhood overweight.

OBJECTIVE. Our goal was to evaluate the relative effectiveness of parenting-skills training as a key strategy for the treatment of overweight children.

DESIGN. The design consisted of an assessor-blinded, randomized, controlled trial involving 111 (64% female) overweight, prepubertal children 6 to 9 years of age randomly assigned to parenting-skills training plus intensive lifestyle education, parenting-skills training alone, or a 12-month wait-listed control. Height, BMI, and waist-circumference z score and metabolic profile were assessed at baseline, 6 months, and 12 months (intention to treat).

RESULTS. After 12 months, the BMI z score was reduced by ~10% with parenting-skills training plus intensive lifestyle education versus ~5% with parenting-skills training alone or wait-listing for intervention. Waist-circumference z score fell over 12 months in both intervention groups but not in the control group. There was a significant gender effect, with greater reduction in BMI and waist-circumference z scores in boys compared with girls.

CONCLUSION. Parenting-skills training combined with promoting a healthy family lifestyle may be an effective approach to weight management in prepubertal children, particularly boys. Future studies should be powered to allow gender subanalysis.
TREATMENT OF CHILDHOOD overweight is an important part of the multilevel response to the obesity epidemic.1,2 Energy intake moderation, increased physical activity, reduced sedentary activity, behavior modification, and family involvement are the assumed cornerstones of child weight management.3-5 However, the most effective and age-appropriate ways to implement these interventions remain unclear.

Familial clustering of overweight,6 the shared family environment,7 and the influence of parents on children’s lifestyle patterns and food choice8 highlight the vital role of family in the management of childhood overweight.9 The developmental stage of children is also an important consideration in developing family-based involvement and behavior-modification strategies.5,9-13 Results from an Israeli study of 50 6- to 11-year-olds suggest that when parents take sole responsibility for managing child overweight, the prevalence of overweight at 8 years’ follow-up is approximately half that found when the child is required to implement lifestyle changes as part of the treatment program.12

The evidence supporting parent-led family lifestyle management for treatment of overweight in young children has not been replicated in other populations. In addition, effective strategies to facilitate parents initiating and maintaining recommended eating and activity behaviors have not been explored. Although parents play a significant part in shaping and influencing child behavior, they rarely receive support or training for this role.14 Parents perceive they possess appropriate nutrition knowledge and are able to assess the dietary adequacy of their child’s diets.15,16 If this is so, then a focus on behavior modification rather than on nutrition education may be appropriate. Parenting-skills training may be an effective age-appropriate child behavior-modification strategy applicable to the management of overweight in young children.

The aim of this study was to evaluate the relative effectiveness of parenting-skills training as a key strategy for the treatment of overweight children. It tests the hypothesis that prepubertal children whose parents participate in a family-focused child weight-management program, comprising parenting-skills training and intensive lifestyle education, will have BMI and waist-circumference z scores and metabolic profiles after 12 months that are (1) improved when compared with children who are wait-listed for intervention for 12 months and (2) no different from children whose parents participate in a program that focused on parenting-skills training alone (ie, without intensive lifestyle education).

METHODS

Participants

Families were recruited between July 2002 and August 2003 predominantly via media publicity and school newsletters. Inclusion criteria were child age 6 to 9 years, overweight (according to the International Obesity Task Force definition17,18), and Tanner stage 119 with a caregiver willing to attend sessions and able to read and understand English. Exclusion criteria were BMI z score > 3.5, diagnosed with a syndromal cause of obesity, using medications that influence weight gain or loss, a diagnosis of physical or developmental disability or chronic illness, and a sibling enrolled in the study. Eligibility was initially assessed via telephone interview and confirmed by a medical practitioner. Parent informed written consent and child assent were obtained.

Study Design

A single-blinded, randomized, controlled trial (Australian Clinical Trial Register 00001103 [www.actr.org.au]) was used to determine the effectiveness of 2 child weight-management interventions, namely parenting-skills training with intensive lifestyle education (P + DA) and parenting-skills training alone (P). These interventions were compared with each other and with a control group wait-listed for intervention for 12 months (WLC). Parents in the P and WLC groups both received a general “healthy-lifestyle” pamphlet. The study was conducted at 2 metropolitan teaching hospitals in Adelaide, South Australia, and was approved by the Flinders Clinical Research Ethics and the Women’s and Children’s Hospital’s ethics committees. The design, conduct, and reporting of this study followed the guidance outlined in the Consolidated Standards of Reporting Trials (CONSORT) statement.20

Randomization schedules were computer generated using a 3-block design stratified for gender and site of recruitment. Individual group allocations were sealed in opaque envelopes, with the next envelope opened on a child’s completion of baseline measurements. Researchers involved in recruitment, participant allocation and intervention delivery (Dr Golley) or data collection (Dr Magarey) were not involved in the randomization process.

All intervention sessions were conducted by the same dietitian (Dr Golley) who had developed the lifestyle education component and undertaken accredited training for the parenting component. The mode of both interventions was “parent only,” with parents having sole responsibility for attending program sessions and implementing family lifestyle change. Children did not attend any education sessions, and families were encouraged to implement change at the family, not child, level.

Intervention Descriptions

P Group

Parenting-skills training was used to facilitate and support parents to undertake family lifestyle change. Par-
ents participated in the Positive, Parenting Program (Triple P, Families International, University of Queensland/Health Department Western Australia, 2000, www.triplep.net), a standardized and evaluated general parenting program.21 Triple P is based on child development theory and social learning principles and aims to promote parental competence to manage their child’s behavior.14 The program consisted of 4 weekly 2-hour group sessions followed by 4 weekly, then 3 monthly 15- to 20-minute individual telephone sessions. Standard Triple P resource materials were used with program examples adapted to reflect dietary and activity behaviors.22 Application of Triple P to eating and activity behaviors was supported by provision of a general healthy-lifestyle pamphlet.

P+DA Group
Parents in the P+DA arm completed the Triple P program as described for the P group above. Parents in the P+DA group participated in an additional 7 intensive lifestyle support group sessions.22 These sessions commenced after completion of the 4 weekly parenting sessions, every 2 weeks at first, then monthly. These sessions focused on lifestyle knowledge and skills including the following: family-focused healthy eating with specific core food serve recommendations,23,24 monitoring, label reading, snacks, modifying recipes, being active in a variety of ways, roles and responsibilities around eating, managing appetite, self-esteem, and teasing.22

While parents attended the lifestyle sessions, children in the P+DA group attended structured, supervised activity sessions developed by physical activity experts. The sessions consisted of fun, noncompetitive games designed around aerobic activity and development of fundamental motor skills. Sessions were designed as play rather than exercise and were diversional rather than interventional. The activities required minimal equipment and were deliverable by nonexpert staff and easily replicated at home.

WLC Group
At the time of group allocation, the WLC group received the same general healthy-lifestyle pamphlet as the parenting alone group. During the 12-month wait-listed period, the WLC group was contacted by telephone 3 to 4 times for 5 minutes as a retention strategy. Researcher contact with the WLC families was minimized to avoid the potential placebo effect of therapist contact.

Measurements
Baseline measurements occurred before randomization with outcome measures assessed at program completion (6 months) for participants in intervention groups and at 12 months after baseline for all participants. Data collection was performed by the same trained assessor who was blinded to participant group allocation.

At baseline an 18-item demographic questionnaire was completed and included parent characteristics (gender, age, ethnicity, relationship to child), family structure (marital status, number of children in the family), and postcode.25 Socioeconomic status was assessed by using the Australian Socio Economic Index for Areas (SEIFA) postcode index of relative advantage. The index is standardized to have a mean of 1000 ± 100, with 95% of index scores falling between 800 and 1200 (high scores indicate high income, skilled labor).26

Anthropometry
The primary study outcome was BMI z score. Height and weight were measured with participants lightly clothed and without shoes. Height to the nearest 1.0 mm was measured with a Trumeter stadiometer (Trumeter, Manchester, United Kingdom), and weight was measured to the nearest 0.1 kg with SECA electronic scales (SECA, Hamburg, Germany). In the absence of national Australian data, BMI was calculated and converted to a BMI z score by using United Kingdom reference data provided as a computer program (Child Growth Foundation, London, United Kingdom).17 Waist-circumference measurement was recorded to the nearest millimeter, midway between the tenth rib and the iliac crest, with participants in a standing position, using a nonelastic flexible tape, and converted to a z score by using United Kingdom reference data.27,28 For categorical analysis, participants were classified as nonoverweight, overweight, or obese using the International Obesity Task Force definition.18 Parental height and weight were either assessor-measured or self-measured (5%–13% of mothers and 71%–78% of fathers self-measured at baseline and follow-up measurements), and BMI was calculated. Parents’ weight status was classified using the World Health Organization definition, with BMI ≥25kg/m² overweight and ≥30kg/m² obese.1

Metabolic Health Outcomes
Blood pressure was measured on the right arm by using a Dinamap automated blood pressure monitor (GE Healthcare, Giles, Buckinghamshire, United Kingdom). A variety of cuff sizes were used to ensure appropriate fit for arm circumference.29 A single measurement was taken after supine rest for 10 minutes after collection of the blood sample and anthropometric measures.29

Fasting glucose, total cholesterol, high-density lipoprotein cholesterol, and triacylglycerol levels were analyzed within 4 hours of collection by standard automotive techniques using a Synchron CX5 Pro analyzer (Beckman Coulter Inc, Fullerton, CA) in the Clinical Diagnostic Laboratory, Women’s and Children’s Hospital, Adelaide, South Australia. Low-density lipoprotein cholesterol was calculated by using the Friedewald equation: cholesterol – (high-density lipoprotein + triacylglycerol/2.2).30 Immediately after sample collection, se-
rum was drawn off and stored at −70°C until fasting serum insulin was measured in batched samples in the Endocrine Diagnostic Laboratory, Royal Prince Alfred Hospital, Sydney, Australia by radio-immunoassay using the Linco human insulin-specific assay kit (Linco Research Inc, St Charles, MO).

Program Evaluation
Parent satisfaction with the intervention programs was assessed by using a validated, anonymous 16-item questionnaire adapted from the one usually used as part of the Triple P program.21 Adaptations included tailoring questions to a child weight rather than a general child behavior management program. Additional questions about perceived barriers to program attendance and implementation were included. Parent attendance at sessions was recorded.

Data Analysis
Sample size calculation was based on a fall in BMI z score reflecting a weight gain of only 50% of that expected over 12 months with normal growth. A sample size of 28 per group was estimated to have 80% power to detect a 12-month fall in mean BMI z score from a baseline of 0.26 ± 0.49, assuming no change in the control group, at a 2-sided significance level of .05. To account for a dropout rate of up to one third (commonly 20%–50% in child weight-management studies), 42 children per study group were sought (N = 126).

Analyses were performed using SPSS for Windows version 11.5 (SPSS Inc, Chicago, IL). Where the distribution of variables is normal, data are expressed as mean ± SD and proportions. Potential covariates were measured at baseline (weight status, growth potential, gender, parental weight status, ethnicity, age, and socioeconomic status), and differences by study group and/or gender were explored by using 1- or 2-way analysis of variance or χ². Baseline differences between those who did and did not attend follow-up measurements were explored by using separate variance t tests.

Intention-to-treat analysis was performed, with all participants included in the analysis according to original group allocation, and follow-up was maximized regardless of program attendance.31 Where variables were normally distributed and had equality of variance of residuals, a linear mixed model (SPSS MIXED) including time (as repeated factor), group, and their interaction, with Bonferroni correction for posthoc multiple comparison, was used to determine whether there was a significant time by group effect between baseline, 6, and 12 months.32,33 Where group by time interactions were non-significant, average intervention effects of the follow-up period were estimated and tested by using the Bonferroni method for posthoc analysis. Secondary analyses were undertaken (1) with gender as a factor and (2) by “per-protocol analysis” including families who attended ≥75% of the program sessions.

RESULTS
Participant characteristics at baseline are shown in Table 1. Sixty-four percent were female, and a majority were 8 years of age or older (75 of 111) and obese (82 of 111). Seventy-two percent of children came from dual-parent families, with 98% of the parents who completed the baseline demographic information being of white ancestry. The mean ± SD for the SEIFA index of relative advantage was 997 ± 73 (South Australian mean: 960). Thirty-four percent of parents were classified as overweight and 44% as obese. At baseline, there were no significant differences for any child or family characteristics by study group.

Figure 1 shows the flow of participants through the study from recruitment to 12-month measurements. There were no significant differences in socioeconomic status (SEIFA indices) between children who enrolled in the study and the 151 who were screened but did not

| TABLE 1 | Anthropometric Measurements (Mean ± SD) for 6- to 9-Year-Old Prepubertal Children at Study Baseline |
|----------------|-------------------------------------------------|-------------------------------------------------|-----------------------------|
|                | All (n = 111)                                    | Boys (n = 41)                                    | Girls (n = 70)               | p value          |
| Age, y         | 8.2 ± 1.1                                       | 8.6 ± 1.0                                       | 7.9 ± 1.2                   |
| Height, cm     | 136.3 ± 8.3                                     | 140.2 ± 7.9                                     | 134.2 ± 7.8                 | <.001            |
| Height z scoreb | 1.25 ± 0.91                                     | 1.37 ± 1.03                                     | 1.17 ± 0.81                 | .01              |
| Weight, kg     | 45.6 ± 9.0                                      | 48.7 ± 10.1                                     | 43.8 ± 7.8                  | .01              |
| BMI, kg/m²b   | 24.3 ± 2.6                                      | 24.5 ± 2.8                                      | 24.1 ± 2.5                  | .41              |
| BMI, z scorec | 2.75 ± 0.52                                     | 2.84 ± 0.43                                     | 2.70 ± 0.56                 | .19              |
| Waist circumference, cm | 77.3 ± 7.3                                     | 80.0 ± 7.5                                      | 75.8 ± 6.8                  | <.01             |
| Waist circumference, z scorec | 3.20 ± 0.65                                     | 3.53 ± 0.67                                     | 3.02 ± 0.57                 | <.001            |

*Independent t test for differences by gender.

b Values were calculated by comparing participant heights against the Centers for Disease Control and Prevention 2000 reference population (Centers for Disease Control and Prevention, Hyattsville, MD: National Center for Health; 2000).41

enroll ($P > .05$). Those who did not attend 12-month measurements were older ($8.7 \pm 0.9$ vs $8.2 \pm 1.2$ years; $P = .04$) and had a higher BMI $z$ score ($2.96 \pm 0.44$ vs $2.71 \pm 0.53$; $P = .04$) and waist-circumference $z$ score ($3.54 \pm 0.64$ vs $3.14 \pm 0.64$; $P = .04$) at baseline than those who did attend.

**Growth, Overall, and Truncal Adiposity**

Height increased in study participants by $6.5 \pm 1.3$ cm between baseline and 12 months. Height $z$ score for all study participants was $1.2 \pm 0.9$ at baseline and $1.3 \pm 0.9$ at 12 months (linear mixed model group by time: $P = .39$), indicating that the growth of intervention children was similar to that of children wait-listed for intervention for 12 months. Over 12 months, the primary study outcome, BMI $z$ score, reduced by 9% (range: $-85\%$ to $18\%$) in the P+DA group, 6% (range: $-48\%$ to $49\%$) in the P group, and 5% (range: $-78\%$ to $16\%$) in the WLC group (linear mixed model, group by time, $P = .76$; Table 2). Fortye-five percent of children in the WLC group increased their BMI $z$ score over 12 months, compared with 19% and 24% in the P+DA and P groups, respectively ($P = .03$).

There was a significant group by time (with gender) interaction for the primary outcome, BMI $z$ score ($P = .04$; Fig 2). Posthoc analysis showed no significant differences between study groups at any time point for boys or girls (all $P > .05$). However, boys in both intervention groups had significantly lower BMI $z$ scores at 6 and 12 months compared with baseline (no change for boys in the WLC group; Fig 2). For girls, the only significant time change was a reduction in BMI $z$ score in the WLC group ($P = .02$). There was no association between change in BMI $z$ score from baseline to 12 months and indicators of socioeconomic status (all SEIFA indices, $P > .05$).

There was a significant group by time interaction for waist-circumference $z$ score ($P = .03$; Table 2). Posthoc analysis showed no significant differences between
study groups at any time point. However, waist-circumference $z$ score was significantly lower at 12 months compared with baseline in the P+DA and P intervention groups but not the WLC group (Table 2) and was also significantly lower at 12 vs 6 months for the P+DA group (Table 2). The pattern of change by gender mirrored that described for BMI $z$ score. For both BMI and waist-circumference $z$ score, results did not change when analysis was conducted per protocol by using only participants with complete data.

### Metabolic Health Outcomes

Table 3 shows the metabolic variables for all participants at baseline and 12 months. There were no differences

<table>
<thead>
<tr>
<th>TABLE 2 Anthropic Outcomes (Mean ± SD) for Study Participants According to Study Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BMI $z$ Score</strong></td>
</tr>
<tr>
<td><strong>Baseline</strong></td>
</tr>
<tr>
<td>(n = 111)</td>
</tr>
<tr>
<td>P+DA</td>
</tr>
<tr>
<td>P</td>
</tr>
<tr>
<td>WLC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Waist Circumference $z$ Score</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
</tr>
<tr>
<td>(n = 111)</td>
</tr>
<tr>
<td>P+DA</td>
</tr>
<tr>
<td>P</td>
</tr>
<tr>
<td>WLC</td>
</tr>
</tbody>
</table>

--- indicates no data collected.

* Linear mixed model: $P = .76$ group by time.

* Linear mixed model: $P = .03$ group by time; posthoc analysis (Bonferroni method): $P < .01$ for P+DA at 12 months versus baseline, $P = .01$ for P+DA at 12 vs 6 months, and $P = .05$ for 12 months versus baseline; for all other comparisons, $P > .05$.

* Mean differences between 12 months and baseline.

<table>
<thead>
<tr>
<th>TABLE 3 Total Cholesterol and Fractions, Triacylglycerol, Blood Pressure, Glucose, and Insulin (Mean ± SD) for 6- to 9-Year-Old Prepubertal Children at Baseline and 12 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
</tr>
<tr>
<td>(n = 11)</td>
</tr>
<tr>
<td>Total cholesterol</td>
</tr>
<tr>
<td>Low-density lipoprotein cholesterol$^b$</td>
</tr>
<tr>
<td>High-density lipoprotein cholesterol$^b$</td>
</tr>
<tr>
<td>Triacylglycerol</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
</tr>
<tr>
<td>Glucose</td>
</tr>
<tr>
<td>Insulin</td>
</tr>
</tbody>
</table>

* Linear mixed model: group by time interaction.

* Linear mixed model: main effect of time ($P < .05$).
between study groups at baseline or 12 months. Diastolic blood pressure was significantly reduced at 6 months but not at 12 months when compared with baseline.

Program Evaluation

Program attendance did not differ between the 2 intervention groups with 18 of 38 and 19 of 37 parents from the P+DA and P groups, respectively, attending more than three quarters of the program sessions. As part of the anonymous satisfaction questionnaire (71% response rate: 26 P+DA, 10 P), parents were asked to circle factors that prevented them attending intervention sessions. Family or work commitments, family illness, and perceived lack of time were more frequently indicated as barriers to intervention attendance (21 of 36), rather than program-related barriers (eg, session timing or frequency, transport difficulty, program not meeting needs, 9 of 36). Seven parents in both intervention programs sought other assistance elsewhere regarding child weight management during the study.

All 36 respondents in both intervention groups rated the quality of the service provided during the interventions as good to excellent. Thirteen of 26 P+DA parents and 8 of 10 P parents reported they “generally to definitely” received the type of help they wanted in managing their child’s weight. All parents in the P group and 22 of 26 in the P+DA group were “satisfied to very satisfied” with the amount of help received during the study. All parents in the P group and 24 of 26 in the P+DA group responded that the study had “helped somewhat to helped a great deal” to make changes to family lifestyle. Twenty of 26 P+DA respondents and 6 of 10 P group respondents said they would repeat the program if they were seeking assistance in managing child overweight again. The parenting-skills training resources (group parenting sessions: 14 P+DA, 7 P; parenting telephone sessions: 12 P+DA, 7 P; parenting manual: 10 P+DA, 6 P) were more commonly reported as being useful than the lifestyle education resources (lifestyle sessions: 5 P+DA, not applicable for P; lifestyle written material: 5 P+DA, 1 P).

Discussion

This study examined the effectiveness of parenting-skills training, with or without intensive lifestyle education, as part of parent-led, family-focused weight management of 6- to 9-year-olds. The key study finding was that all 3 groups had a significant reduction in BMI z score over 12 months. Although there was no statistical significance between groups BMI z score decreased over 12 months in double the number of children in the P+DA group (45%) compared with the P intervention or intervention wait-listing (24% and 19%, respectively). Waist-circumference z score fell significantly over 12 months in both intervention groups but not in the control group. Importantly, when time and group analysis was adjusted for gender boys in the intervention groups had significant reductions in both BMI and waist-circumference z scores, which were not observed for girls or the wait-listed controls. Because a gender effect was unexpected, the study was not powered to allow gender subanalysis.

In terms of broader health outcomes, there was no significant change in metabolic profile in any study group. The average values for metabolic indicators at baseline were within the reference range for children, thus limiting the need for improvement over time.30,34 The interventions had no adverse impact on linear growth.

Only 1 previous study examined differences in intervention response by gender.35 As was found in our study, boys showed better reductions in percent overweight than girls in a family behavioral program that focused on increasing activity and decreasing sedentary activity.35 In that study, gender did not independently predict reductions in percent overweight, and the authors proposed that the effect of gender was mediated by the higher motivation and intervention compliance observed with families of boys.35 In our study, the higher degree of truncal adiposity in boys versus girls at baseline may be a factor in the gender differences in intervention response.36

Recent data suggesting that truncal adiposity in Australian children is increasing more rapidly for girls than boys may provide an additional explanation for the gender differences.37 There may be gender-based differences in the environmental influences promoting overweight, making it harder in girls to reverse this trend. Our results highlight the need to always stratify for gender when allocating participants and to ensure that sample size calculations allow gender subanalyses.

The weight status changes from this study, although clinically relevant, were modest compared with previous studies. Child-focused studies using the same outcomes as the present study have achieved between 0.6 and 1.3 unit reductions in BMI z score compared with ~0.25 in this study.36–40 However, analysis in these studies were not performed by intention to treat and may have overestimated intervention effectiveness.33 The only other study that evaluated parent-led child weight management had a 15-point reduction in percent overweight (using percent above the 50th centile for weight) compared with an 8-point reduction with a child-focused intervention.41 Comparison of these results to the present study is limited by the difference in outcome expression and the length of intervention (12 vs 6 months in our study). This is the first time that waist-circumference z scores have been reported in a child weight loss study.

The parenting-skills training only group was included to explore the commonly espoused idea that parents have sufficient lifestyle knowledge but lack active parenting/behavior-modification skills to support weight
management. Although it seems that intensive lifestyle education may enhance parents’ ability to implement parent-led child behavior modification, an intensive lifestyle education alone group would be required to fully evaluate the relative role of parenting-skills training and lifestyle education or synergy between these treatment components. The potential resource savings of parenting-skills training alone, if effective, also has important health cost implications.

Our study addressed 5 of the major design limitations of child weight-management studies highlighted in a recent Cochrane review including (1) using recruitment and retention strategies to achieve the calculated sample size and <20% drop out at 12 months, (2) blinded allocation and outcome assessment to minimize measurement bias, (3) intervention delivery using standard protocols and a single, trained facilitator to limit site bias and enhance internal study validity, (4) health outcome protocols and a single, trained facilitator to limit site bias, (5) intervention delivery using standard allocation and outcome assessment to minimize measurement bias, (3) intervention delivery using standard allocation and outcome assessment to minimize measurement bias, (4) health outcome protocols and a single, trained facilitator to limit site bias, (5) primary analysis performed by intention to treat to properly assess intervention effectiveness. Furthermore, the narrow age range ensured age-appropriate intervention.

Compared with previous studies, considerable attention was paid to defining effectiveness and based on broad criteria, the results suggest potential for overall intervention effectiveness. However, study power, intervention adherence, and dilution of effect size with intention-to-treat analysis may be factors in the failure of these trends to convert to statistical significance and the moderate results compared with previous studies.

Sample size based on the primary outcome assumed there would be no change in adiposity in the control group. Hence, the unanticipated reduction in BMI z score in the control group produces the potential for type II error. The apparent motivation of the control group limits generalizability; however, in the current obesity epidemic environment and associated media coverage, such bias may be difficult to avoid.

CONCLUSIONS
A family-focused intervention using parenting-skills training and promoting a healthy family lifestyle may be an effective approach to weight management in prepubertal children but with a clear gender effect. Both parenting-skills training and lifestyle education are potentially important components. This approach addresses family and parental factors influencing children's eating and activity behaviors and achieves a moderate reduction in adiposity after 12 months. Future studies should be powered for adiposity reductions in control groups, primary and secondary outcome analysis and gender subanaylisis. Finally longer-term follow-up to assess effectiveness through to early adolescent development is essential in the context of global obesity.

ACKNOWLEDGMENTS
This research was funded by the Australian Health Management Group Assistance to Health and Medical Research Fund. Dr Golley was supported by Australian National Health and Medical Research Council Postgraduate Research Scholarship 229978.

We thank the children and families for participating in the study, including blood sample collection. Many thanks go to the following for assistance: Professor Kevin Norton (School of Health, University of Adelaide) for developing the physical activity component; staff from the paediatric unit at Flinders Medical Centre for provision of space and screening participants; Ms Sarah Garnett (Institute of Endocrinology, Children’s Hospital at Westmead) and Ms Margie Grucu (laboratory manager, James Fairfax Institute Pediatric Nutrition, Children’s Hospital at Westmead) for calculation of waist-circumference z scores; Ms M. Howard, Dr D. Thomas, and the Division of Laboratory Medicine (Women’s and Children’s Hospital, Adelaide); Diagnostic Laboratory, Department of Endocrinology Royal Prince Alfred Hospital, Sydney; and Ms Kylie Lange and Ms Fiona Curnow for statistical advice and support.

REFERENCES
12. Golan M, Crow S. Targeting parents exclusively in the treat-


Work-Related Hazards and Workplace Safety of US Adolescents Employed in the Retail and Service Sectors

Carol W. Runyan, PhD, Michael Schulman, PhD, Janet Dal Santo, DrPH, J. Michael Bowling, PhD, Robert Agans, PhD, Myduc Ta, MPH

OBJECTIVE. Our goal was to examine the hazard exposures, work experiences, and workplace safety training of adolescents employed in retail and service jobs in the United States.

METHODS. This was a cross-sectional telephone survey among working adolescents, 14 to 18 years old, in the continental United States. Data were collected in 2003. Survey items measured self-reported hazard exposures, training, and supervision experiences of working adolescents.

RESULTS. Teens reported working an average of 16.2 hours per week during the school year, including working an average of 2.9 times per week after 7 PM on school nights and 2.6 nights per week after 9 PM. Thirty-seven percent of those under age 16 reported working after 7 PM on a school night, indicating employer violation of federal law. Teens typically perform multiple kinds of tasks in a given job. Higher proportions of females than males are involved in cash handling (84% vs 61%), whereas males are more likely than females to be involved in physically challenging tasks, such as lifting heavy objects (57% vs 22%) or working at heights (35% vs 17%). Despite federal regulations prohibiting teens under 18 from using certain types of dangerous equipment (eg, slicers, dough mixers, box crushers, paper balers) or serving or selling alcohol in places where it is consumed, 52% of males and 43% of females reported having performed ≥1 prohibited task. Although more males reported receiving safety training, they were also more likely to report working without supervision than their female counterparts.

CONCLUSIONS. Teens are exposed to multiple hazards, use dangerous equipment despite federal prohibitions, and work long hours during the school week. They also lack consistent training and adult supervision on the job. It is important for adolescent medicine practitioners to become involved in prevention efforts through both anticipatory guidance and policy advocacy.
In 2000, the estimated labor force participation for 16- to 17-year-olds was 68.5% (75.3% for males and 62.1% for females), resulting in >6 million employed adolescents in the United States. Many teens, however, begin working before their 16th birthdays. Once adolescents enter the labor market, they usually continue working, although they change jobs frequently. Most working adolescents are employed during both the school year and the summer, with the proportions working during both periods estimated at 62% for 16- to 19-year-olds. Estimates suggest that by the time they graduate from high school, most teenagers have had a job. The largest proportion (62%) of adolescents work in retail trades, of which half are eating and drinking establishments, whereas the second largest number (25%) work in the service industry during school months. In the summer months, retail and services industries account for 50% and 30% of youth employment, respectively.

Despite its benefits, employment can have serious negative consequences for adolescents. Several studies document the magnitude of fatal injuries to young workers. The Bureau of Labor Statistics reports fatality counts of work injuries among workers 17 years old and younger to be fairly stable between 1992 and 2000, with an annual average of 68 fatalities.

Converting actual hours worked to full-time equivalency (because youth work part-time or for limited time periods) results in estimated fatality rates per 100,000 workers of 5.1 for 15-year-olds, 3.4 for 16-year-olds, and 3.7 for 17-year-olds. Although these rates of fatal work-related injuries to teens are the same or slightly less than the rate for adults (5.1 per 100,000), the number of injuries per hour worked may actually be greater for youth.

Injuries to youth in the retail trade and services sectors together accounted for 26% of the fatalities among workers 17 years old and under during the 1998–2002 period. For this period, 13% (40) of the 304 deaths to young workers were in the retail trades, and service industries accounted for another 13% of the documented young worker fatalities. Homicides associated with robberies were reported to be the cause of between one fourth and one half of all youth fatalities in the retail trade.

With regard to nonfatal work injuries, Layne et al. estimated that 64 100 youths between the ages of 14 and 17 were treated in emergency departments for occupational injuries in 1992 and National Institute for Occupational Safety and Health estimates that nearly 200,000 adolescents are injured at work every year. A wide variety of data sources have been used to characterize nonfatal adolescent work-related injuries. Several studies used workers’ compensation data to determine the incidence of nonfatal injuries to teen workers, whereas others used first report of injury records or focus groups and surveys of adolescents. Data from the Survey of Occupational Injuries and Illnesses, an annual survey of employers’ Occupational Safety and Health Administration (OSHA) recordkeeping logs conducted by the Bureau of Labor Statistics, reported 9000 cases of workers younger than age 17 with injuries and illnesses that resulted in days away from work in 2003. Approximately 25% of these cases were in retail trade, and 38% were in food services and drinking places.

This study describes the self-reported work experiences of a national sample of employed adolescents, ages 14 to 17, working in the retail and service sectors. We documented employee hours of work, the tasks the youth perform, and the hazards that they face on the job, examining variations by gender and age. In addition, we explored their reported experiences with being trained and supervised. As such, this is the first national study, to our knowledge, to document the work-related hazards to which teens are exposed.

METHODS

Overview

This project used a nationally representative cross-sectional telephone survey to elicit information about work experiences from a sample of 14- to 17-year-old working teens.

Instrument Development

To develop our survey instrument, we conducted focus groups with teen workers to learn how they described workplace hazards, reviewed teen worker injury reports obtained from the state of Massachusetts (Leticia Davis, ScD, verbal communication, 2002), and borrowed items from national surveys and from previous surveys of teens in North Carolina and in selected other sites. We then sent the draft instrument for review to 7 national experts familiar with teen labor issues, making revisions before pretesting to clarify confusing questions and reduce the length of survey administration to ~25 minutes. We pilot-tested the instrument with 54 teens incorporating final modifications in preparation for interviewer training and final survey administration by the Survey Research Unit of the Department of Biostatistics at the University of North Carolina School of Public Health.

The Human Subjects Committee of the University of North Carolina School of Public Health reviewed and approved the instrument, as well as consent, assent, and interview procedures.

The final survey instrument, which was programmed into a computer-assisted system using Blaise, consisted of 50 items, including categories addressing: business type and job tasks, work hours, and hazard exposures.
Sampling
A probability sample of households with telephone line access in the continental United States was chosen for this study. The sample was selected by using a dual-frame approach in which 10% of the sample was selected from a list-assisted random-digit dialing frame of households with telephones, and the remainder was chosen from a complementary (ie, nonoverlapping) targeted-list frame consisting of directory-listed residential telephone numbers for which recent demographic information about the household was available.

Screening and Eligibility
Once they reached a residential household with a working telephone number, interviewers determined the eligibility of the household by talking with a parent or guardian. A household was eligible to participate in the study if it had a teen between the ages of 14 and 18 who had held a job not supervised by a parent/guardian for at least a 2-month period within the preceding 12 months. In addition, the teen needed to be ≥14 years old but younger than 18 at the time he or she worked. If there was >1 eligible teen in the household, 1 was selected randomly to participate. We asked the parent to indicate all the eligible teens. Depending on how many eligible teens were in the household, we generated a random number from that number of possible numbers (ie, if 3 teens were present, we selected randomly number 1, 2, or 3). If “1” was generated, we selected the youngest eligible child. If “2” was generated, we selected the second youngest eligible child, and so on.

The parent or guardian gave consent before data collection could begin and then was typically interviewed for the parent part of the study (C.W.R., M.S., J.D., J.M.B., and R.A., unpublished data, 2003) before we asked to speak with the adolescent. The parent and the selected teen were each read statements explaining the study and confidentiality provisions and were asked to agree verbally to participate. Non–English-speaking households (n = 418), as well as households without telephones, were excluded from the study population. As part of the parent interview, we asked the parent to identify the most recent job meeting the eligibility criteria. This job selection was confirmed with the teen during the teen interview so that the teen’s responses focused on 1 specific job during the entire interview.

Data Collection
Interviews took place 7 days a week between the months of February and September 2003. Selected numbers were called until a minimum of 10 unsuccessful attempts were made with at least 1 weekend call, 1 evening call, and 1 daytime call over a several-week period.

Data Management and Data Analysis
Raw sample weights were based on original probabilities of choosing households from the a national frame of telephone numbers, the total number of telephone lines reaching the household, and in the case of child and parent/child pairs, the number of eligible teens in the household.Trimming techniques based on Potter were used to reduce the overall variability of raw weights. Weight variability inflates standard error of estimates, thereby decreasing the precision of point estimates (proportions or means) and reducing power to make comparisons. The trimming process reduced extreme weights, redistributing them among others so that the sum of the adjusted weights remained constant. Trimming procedures reduced the impact of weight variability on variance of estimates by ~80%. Poststratification adjustments were used to better align multivariate sample distributions on key demographic variables, with population distributions based on national sources. Sample proportions were adjusted to national estimates provided by the 2002 Current Population Survey. Race of household head and household income were used as the poststratification variables.

Analyses showed here include actual frequencies of responses to key items and comparisons within age and gender strata. Percentages, 95% confidence limits (CLs), and means calculations rely on sampling weights.

RESULTS
Response Rate and Final Study Population
Telephone interviews were completed with 928 teenage workers. Using the American Association for Public Opinion Research Standard Definitions, the range in response rates was 51% to 64%. The low-end response rate assumes that the same proportion of unknown eligibility households were eligible to participate as the proportion eligible with known status, and the high-end response rate (64%) assumes that all households contacted for which no eligibility information was available were considered not eligible for the participation in the survey.

Because 93% of the working teens surveyed in this study indicated that their most recent job meeting our eligibility criteria was in retail and service trades, the data presented in this article are restricted to that subset, a study group of 866 teen workers. This study population was 48% male, 85% white, and 3% Hispanic. Thirteen percent of the respondents reported on work experiences at jobs held over the previous year while they were under age 16, and the remaining 21% described experiences working as a 16-year-old and 67% as a 17-year-old.

Work Hours and Schedule
Teens worked at their referent job (ie, the job they had worked at the most during the previous 12 months) for
a median total period of 9 months (interquartile range: 11.0). During the school year, teens reported working a mean of 16.2 hours per week, with 82% of the working teens reporting having worked after 7 PM on a school night, 52% indicating they had worked after 9 PM, and 10% reporting they had worked after 11 PM. The average number of days spent working these hours on school nights also varied, with teens reporting working after 7 PM on average 2.89 nights per week versus working an average of 2.64 and 2.07 nights per week after 9 or 11 PM, respectively (Table 1).

Although not shown in the table, older teens (aged 16–17) were 3 to 4 times more likely to report working in the evening than were teens under age 16. However, 37% of those under age 16 reported having worked past 7 PM on a school night, a practice that is prohibited by the federal law for those under age 16.40 In addition, 16% of these youngest workers (<15 years) indicated they had worked after 9 PM on a school night.

Tasks Performed
Each teen was asked an open-ended question: “What tasks do (did) you do at this job?” Up to 5 responses to the question were coded. These tasks were categorized postinterview into 8 groups: (1) customer interaction and cash handling tasks (eg, cashiers, baggers, hostesses, and sales); (2) food preparation and food handling and serving tasks (eg, cooks, bartenders, waiters, and deli work); (3) cleaning and laundry tasks (eg, busing tables, dishwashing, laundering, and removing trash); (4) child and health care (eg, work in tending to children in day care or elders in nursing homes); (5) recreation, sports, and teaching (eg, coaching, refereeing, teaching sports, camp counseling, life guarding, and preparing golf carts for patrons); (6) driving and delivery (eg, food delivery, steering boats, or driving cars); (7) other manual tasks (eg, loading, stocking shelves, operating equipment, doing yard work, animal care, and repairing items); and (8) other, including office work (eg, filing, typing, and doing work on computers). Some responses were too vague to be classified and were coded as “unknown.”

Overall, 17% of the teens listed just 1 task at their referent job, whereas 24% listed 2, an additional 24% reported 3 tasks, and 34% listed 4 or more tasks. As shown in Table 2, 31% of females and 24% of males reported a primary job task involving customer interaction or cash handling, and 30% of females compared with 20% of males indicated their primary task was preparing, handling, and/or serving food. Although not shown, responses to another question indicated that 84% of females and 61% of males indicated that they had ever been responsible for cash handling in this job, even if it was not their primary task. Males were more likely to identify primary tasks of doing cleaning and laundry work (18% of males vs 8% of females) or manual tasks such as loading trucks, stocking shelves, yard work, or assembling things (17% of males vs 7% of females). Higher proportions of males than females reported their primary task was recreation, sports, or teaching (12% of males vs 8% of females) whereas females were more engaged with child care as the primary task (5%) versus males (3%).

Because we were particularly interested in learning about hazards on the job that might lend themselves to remediation, respondents were asked about tasks specific to retail and service sector employment shown in previous studies to be potentially dangerous.22,41 Items in Table 2 were asked of all respondents, whereas those in Table 3 were asked only of those 395 respondents who worked in grocery or food service establishments. It is important to note that in Table 2, the sample sizes and, therefore, stability of the estimates vary because we asked individual items only to those respondents working in establishments that had the specified type of equipment.

As shown in Table 2, close to 90% of teen workers reported performing cleaning tasks (as distinct from only the primary tasks). Among the more than three quarters

<table>
<thead>
<tr>
<th>TABLE 1 Work Hours During School Year of US Teens Working in Retail and Service Sectors, 2003 (N = 866)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Have you worked when there is school the next day (ever)?</td>
</tr>
<tr>
<td>Work after 7 PM on school night</td>
</tr>
<tr>
<td>Work after 9 PM on school night</td>
</tr>
<tr>
<td>Work after 11 PM on school night</td>
</tr>
<tr>
<td>Frequency of work after specific hours on school nights, No. of times in an average week, mean (range)</td>
</tr>
<tr>
<td>Work after 7 PM</td>
</tr>
<tr>
<td>Work after 9 PM</td>
</tr>
<tr>
<td>Typical No. of hours worked per week when school is in session, mean (range)</td>
</tr>
</tbody>
</table>
TABLE 2  Primary Tasks and Other Selected Experiences Reported by US Working Teens in the Retail and Service Sectors According to Gender, 2003 (N = 866)

<table>
<thead>
<tr>
<th>Working Conditions</th>
<th>Female, % (95% CLs)</th>
<th>Male, % (95% CLs)</th>
<th>Total, % (95% CLs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary tasks (n = 866)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Customer interaction, cash handling</td>
<td>31.4 (24.2, 38.6)</td>
<td>23.6 (17.3, 29.9)</td>
<td>27.7 (22.8, 32.5)</td>
</tr>
<tr>
<td>Food preparation, food handling, serving</td>
<td>30.3 (21.9, 38.6)</td>
<td>20.4 (13.7, 27.1)</td>
<td>25.5 (20.0, 31.0)</td>
</tr>
<tr>
<td>Cleaning and laundry tasks</td>
<td>8.0 (3.8, 12.1)</td>
<td>17.5 (10.1, 24.8)</td>
<td>12.6 (8.3, 16.8)</td>
</tr>
<tr>
<td>Child and health care</td>
<td>5.0 (2.6, 7.4)</td>
<td>2.7 (0.55, 4.8)</td>
<td>3.9 (2.3, 5.5)</td>
</tr>
<tr>
<td>Recreation, sports and teaching</td>
<td>8.1 (5.0, 11.2)</td>
<td>12.4 (8.7, 16.0)</td>
<td>10.2 (7.8, 12.5)</td>
</tr>
<tr>
<td>Driving and delivery</td>
<td>0.8 (0.0, 2.2)</td>
<td>1.2 (0.19, 2.2)</td>
<td>1.0 (0.1, 1.9)</td>
</tr>
<tr>
<td>Other manual tasks (eg, loading, stacking shelves, operating equipment)</td>
<td>7.2 (2.5, 11.9)</td>
<td>16.9 (11.1, 22.8)</td>
<td>11.9 (8.2, 15.7)</td>
</tr>
<tr>
<td>Other, including office work, work with computers</td>
<td>9.2 (5.5, 12.8)</td>
<td>5.3 (0.65, 9.8)</td>
<td>7.3 (4.4, 10.2)</td>
</tr>
<tr>
<td>Other experiences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor vehicle present at workplace (ie, company motor vehicles that employees use as part of their job), females (n = 90); males (n = 118)</td>
<td>22.5 (4.0, 4.0.9)</td>
<td>40.1 (24.3, 55.9)</td>
<td>32.1 (20.3, 44.0)</td>
</tr>
<tr>
<td>Lawnmowers present at workplace, females (n = 76); males (n = 115)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you operated a lawn mower?</td>
<td>14.0 (0.0, 34.3)</td>
<td>19.9 (9.0, 30.8)</td>
<td>17.2 (6.4, 27.9)</td>
</tr>
<tr>
<td>Power equipment or tools present at workplace, females (n = 162); males (n = 223)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you operated power equipment or tools?</td>
<td>31.0 (18.4, 43.6)</td>
<td>59.4 (49.1, 69.8)</td>
<td>47.4 (39.2, 55.7)</td>
</tr>
<tr>
<td>Forklift or any other power-driven lifting equipment present at workplace, females (n = 57); males (n = 98)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you operated a forklift or other power driven lifting equipment?</td>
<td>6.1 (0.0, 14.6)</td>
<td>44.6 (28.4, 60.8)</td>
<td>28.6 (16.2, 41.0)</td>
</tr>
<tr>
<td>Heavy equipment or machinery such as that used in cleaning, landscaping, construction, or industrial work at present workplace, females (n = 64); males (n = 116)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you operated heavy equipment, machinery for cleaning, landscaping, construction, or industrial work?</td>
<td>3.7 (0.1, 7.4)</td>
<td>26.6 (15.4, 37.8)</td>
<td>17.4 (10.4, 24.4)</td>
</tr>
<tr>
<td>Have you been an outside helper on a motor vehicle? Females (n = 428); males (n = 438)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you worked in high places (&gt;6 ft)?</td>
<td>16.6 (10.2, 22.6)</td>
<td>25.4 (17.9, 31.7)</td>
<td>25.6 (20.8, 30.4)</td>
</tr>
<tr>
<td>worked in high places without fall protection?</td>
<td>28.7 (13.3, 44.2)</td>
<td>49.7 (36.8, 62.5)</td>
<td>42.6 (32.0, 53.1)</td>
</tr>
<tr>
<td>moved or lifted heavy objects (&gt;50 lbs)?</td>
<td>22.1 (15.4, 28.7)</td>
<td>57.0 (49.1, 64.8)</td>
<td>38.8 (33.2, 44.3)</td>
</tr>
</tbody>
</table>

TABLE 3  Percentages of US Working Teens Reporting Specific Activities in Grocery Stores or Food Service According to Gender, 2003 (N = 395)

<table>
<thead>
<tr>
<th>Work Activities</th>
<th>Female (n = 191), % (95% CLs)</th>
<th>Male (n = 204), % (95% CLs)</th>
<th>Total (N = 395), % (95% CLs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used sharp knives</td>
<td>67.7 (55.4, 80.0)</td>
<td>74.9 (64.8, 85.0)</td>
<td>71.2 (63.1, 79.3)</td>
</tr>
<tr>
<td>Used case cutter, box knife, or razor blades</td>
<td>59.3 (47.8, 70.9)</td>
<td>77.3 (67.6, 87.0)</td>
<td>68.2 (60.5, 76.0)</td>
</tr>
<tr>
<td>Used grills or ovens</td>
<td>53.6 (41.7, 65.5)</td>
<td>55.6 (43.8, 67.4)</td>
<td>54.6 (46.2, 62.9)</td>
</tr>
<tr>
<td>Used deep fat fryer</td>
<td>36.0 (24.7, 48.4)</td>
<td>36.7 (25.9, 47.5)</td>
<td>36.4 (28.1, 44.6)</td>
</tr>
<tr>
<td>Used power slicing tool or grinder†</td>
<td>15.8 (8.1, 23.5)</td>
<td>19.3 (10.0, 28.6)</td>
<td>17.5 (11.5, 23.5)</td>
</tr>
<tr>
<td>Sold or served alcohol at places where alcohol is consumed by customers‡</td>
<td>17.0 (5.9, 28.1)</td>
<td>12.7 (3.6, 21.7)</td>
<td>14.9 (7.6, 22.1)</td>
</tr>
<tr>
<td>Used steam table</td>
<td>15.2 (5.9, 24.6)</td>
<td>12.0 (6.6, 17.4)</td>
<td>13.7 (8.1, 19.2)</td>
</tr>
<tr>
<td>Used box crusher†</td>
<td>6.2 (1.5, 10.8)</td>
<td>21.4 (11.2, 31.7)</td>
<td>13.6 (7.8, 19.3)</td>
</tr>
<tr>
<td>Used dough mixing or rolling machine‡</td>
<td>7.2 (2.7, 11.8)</td>
<td>16.3 (7.2, 25.5)</td>
<td>11.7 (6.5, 16.8)</td>
</tr>
<tr>
<td>Used baler or compactor†</td>
<td>4.9 (1.2, 8.6)</td>
<td>17.6 (10.2, 25.0)</td>
<td>11.1 (7.0, 15.2)</td>
</tr>
<tr>
<td>Used food wrapping machine‡</td>
<td>4.1 (1.6, 6.6)</td>
<td>16.5 (7.2, 25.8)</td>
<td>10.1 (5.2, 15.0)</td>
</tr>
<tr>
<td>One or more prohibited tasks reported‡</td>
<td>43.2 (31.1, 55.2)</td>
<td>51.9 (40.0, 63.0)</td>
<td>47.4 (38.9, 55.9)</td>
</tr>
<tr>
<td>Mean number of above hazards reported‡</td>
<td>2.8 (2.4, 3.3)</td>
<td>3.6 (3.1, 4.1)</td>
<td>3.2 (2.9, 3.6)</td>
</tr>
</tbody>
</table>

* These activities are prohibited by federal law for workers <18 years of age. State laws vary and may have additional prohibitions.

of the teens who worked in places with cash registers, the majority were engaged in cash handling, with figures of 84% for females and 61% for males. More than half of the male teens compared with 22% of the females reported having lifted >50 pounds while at work. Similarly, males were more likely than females to work at heights >6 feet (35% vs 17%) and without use of fall protection (50% vs 29%).

Hazard Exposures and Illegal Tasks

Table 3 reports information about the 395 youth working in groceries or food service establishments. Among
these teens, more than two thirds reported using sharp
knives (71%) or using case cutters, box knives, or razor
blades (68%). Over half had used grills or ovens, and
36% had used deep fat fryers. Male workers reported
higher exposures to all these hazards than female work-
ers. Four types of injury hazards listed are explicitly
prohibited by federal law for any worker under age 18:
operation of box crushers, operation of balers or com-
pactors, operation of a power slicing tool or grinder, and
operation of dough mixers. A fifth, selling or serving
alcohol where it is consumed, is also included. More
than half (52%) of all male workers and 43% of female
workers reported having performed ≥1 of these 5 fed-
erally prohibited tasks.

Training and Supervision

Overall, two thirds of the respondents indicated they had
received some type of safety training (Table 4). The
reported content of the training varied. Most of the teens
reported that they were taught about using equipment
safely and how to avoid getting hurt while working.
Fewer teens indicated they had received training related
to violence, with fewer than half reported having been
taught what to do in the event of a robbery. Higher

proportions of females than males report training for
robbery, angry customers, and threats of attack.

We asked several questions about supervisory prac-
tices. More than 60% of teens reported that someone
checked to make sure they were doing their work cor-
rectly at least once a day. In addition, we asked the
question: “In a typical work week while working at
[referent job], how many days do (did) you work some
or part of the day without an adult supervisor (age 21 +
years) at the worksite?” In response, 22% of the females
and 30% of the males reported that in a typical week
they worked without adult supervision ≥1 day. Al-
though not shown in Table 4, 19% of teen workers
reported working without supervision ≥2 days per
week.

Previous studies identify robbery and workplace vio-
lence as major risks associated with working in retail
settings. Therefore, we asked questions about work-
alone during daylight and/or evening hours. Ap-
proximately 10% indicated they had worked alone for ≥
1 day at their referent job during the past year during
daylight hours, and 9% percent indicated that they had
worked alone after dark (for at least half an hour) ≥1
day a week (Table 4). Although not shown in Table 4,

| TABLE 4 | Training and Supervision by Gender, 2003 (N = 866) |
|-----------------|-----------------|-----------------|
|                | Female (n = 428), % (95% CLs) | Male (n = 438), % (95% CLs) | Total (N = 866), % (95% CLs) |
| Training and training content                  |                  |                  |
| Have you received any kind of safety training? | 61.1 (53.3, 68.8) | 72.4 (66.0, 78.8) | 66.5 (61.3, 71.7) |
| Training addressed                              | (n = 258)        | (n = 317)        | (n = 575) |
| How to avoid getting hurt while working         | 86.4 (80.9, 91.9) | 82.2 (73.4, 90.9) | 84.1 (78.9, 89.5) |
| How to use equipment safely                      | 90.0 (85.3, 94.6) | 92.2 (89.7, 96.7) | 91.7 (88.8, 94.5) |
| What to do in case of robbery                    | 44.1 (38.0, 50.4) | 37.7 (28.2, 47.3) | 40.8 (33.7, 47.9) |
| How to deal with angry or drunk customer         | 65.0 (55.5, 74.6) | 56.3 (46.3, 66.4) | 60.5 (53.5, 67.5) |
| How to deal with arguments or fights among coworkers | 64.6 (55.6, 73.6) | 62.3 (53.2, 71.3) | 63.4 (57.0, 69.8) |
| What to do if sexually harassed                  | 66.0 (57.1, 74.9) | 60.8 (51.3, 70.3) | 63.3 (56.7, 69.8) |
| What to do if attacked or threatened             | 63.3 (53.6, 72.9) | 58.6 (48.7, 68.5) | 60.9 (53.9, 67.8) |
| Supervision                                      |                  |                  |
| How many days did you work without adult supervisor at the worksite? | | |
| None                                            | 78.5 (72.7, 84.2) | 70.0 (62.3, 77.7) | 74.4 (69.6, 79.3) |
| ≥1                                              | 21.5 (15.8, 27.3) | 30.0 (22.3, 37.7) | 25.6 (20.7, 30.4) |
| How many days were you the only person at worksite during daylight hours? | | |
| None                                            | 89.1 (84.9, 93.4) | 91.8 (88.4, 95.2) | 90.4 (87.6, 93.2) |
| ≥1                                              | 10.9 (6.6, 15.1)  | 8.2 (4.8, 11.6)   | 9.6 (6.8, 12.4)   |
| How many days were you the only worker at the worksite after dark for at least half an hour? | | |
| None                                            | 92.0 (88.3, 95.7) | 89.1 (82.8, 95.3) | 90.6 (87.0, 94.2) |
| ≥1                                              | 8.0 (4.3, 11.7)   | 10.9 (4.7, 17.2)  | 9.4 (5.8, 13.0)   |
| How often, if ever, has anyone checked to make sure you were doing your work correctly? | | |
| More than once a day                             | 36.2 (27.8, 44.5) | 40.2 (32.1, 48.4) | 38.1 (32.3, 43.9) |
| Once a day                                      | 25.6 (19.0, 32.3) | 25.1 (18.9, 31.4) | 25.4 (20.8, 29.9) |
| At least once a week, but not every day         | 22.1 (15.9, 28.4) | 22.9 (16.0, 29.7) | 22.5 (17.9, 27.1) |
| Less than once a week                            | 11.4 (7.6, 15.1)  | 4.9 (2.7, 7.3)    | 8.3 (6.1, 10.6)   |
| Never                                           | 3.5 (1.0, 6.0)    | 6.5 (1.8, 11.2)   | 4.9 (1.3, 7.5)    |
| It varies                                       | 1.2 (0.0, 2.8)    | 0.3 (0.0, 0.7)    | 0.8 (0.0, 1.5)    |
15% of workers reported having worked alone in 1 or both of these circumstances (daytime or night hours) for \( \geq 1 \) day a week.

DISCUSSION
Overview
Adolescents’ formal employment for wages begins at an early age and, on average, teens in retail and service settings keep their jobs for nearly a year. Teens’ work hours vary between school year periods and school vacations, with teens working nearly twice as many hours per week during school vacations. Older teens and males are exposed to more hazards than younger or female workers. However, by virtue of being more heavily involved in cash handling, females are exposed to the risks associated with robberies.

In addition, many teens work at night on school nights, including some who indicated that they worked after 11 PM on a school night. This suggests the potential for interference with school or sleep, as well as potential for exposure to workplace violence that is more prevalent in the retail and service sectors than in other settings. For teens younger than 16 years of age, working after 7 PM on school nights is illegal and suggests the need for better enforcement of child labor laws. Although teen worker fatalities are less common in the retail and service sectors than in other sectors, nonfatal injuries in these sectors are common. The data from this survey confirm that teens employed in these jobs are exposed to a variety of hazardous tasks, tools, and situations.

Teen work is complex. Workplaces in full compliance with OSHA regulations and standards for ensuring a safe work environment may still place young workers at risk for injury and illness. Federal and state child labor laws are designed to impose additional restrictions to protect teens from hazardous work environments or late work hours. As suggested by our data, many teens are performing tasks that are prohibited by current federal child labor laws. Our results also suggest gaps in both safety training and supervision of working teens because approximately one third of the teens reported not receiving any safety training. Among those who received safety training, training was lacking in some critical areas, such as training on what to do in case of robbery or on how to deal with arguments or fights among coworkers.

Although the efficacy of supervision and training is not well studied, it has been demonstrated that workers working alone are at greater risk of workplace homicide. Plus, it is likely that greater supervision and training in difficult situations that arise in retail and service sector jobs would benefit adolescents who may lack the life experiences and judgment to develop appropriate strategies to deal with complex situations. The fact that so many teens reported working \( \geq 1 \) day a week without adult supervision suggests the potential for serious lapses in safety.

Limitations
Because this is the first national study, to our knowledge, to interview working teens in the United States about their hazard exposures, there are few points of comparison. Rothstein reports that the rates of employment among minority youth are lower than for white teens. Most likely, we have missed immigrant teenagers, particularly those who are undocumented both because of limiting our interview to those parents and teens who can speak English and because of a possible lack of receptivity of these families to agree to be interviewed. In addition, youth working in agriculture or in informal sector jobs (eg, working as a day-laborer or for a relative) were not included.

In addition, we cannot assess the validity of responses that could be biased because of participant recall. The survey relied on self-report of job tasks and hazard exposure. This is a potential source of bias because of problems of recall and lack of accurate measurement of the number of times a particular hazard exposure may have occurred. We attempted to minimize recall problems by asking respondents to focus on a single, referent job throughout the interview. Although we can be sure, we suspect that recall biases are conservative in underrepresenting the true prevalence of hazard exposure.

Finally, although this is the largest study of its kind to date, the sample size is limited, and CLs around some estimates are wide.

Implications for Future Research and Intervention
The literature on teen labor is relatively sparse and derives from several distinct fields, including youth development, public health, sociology, education, and organizational psychology. Little appears in the medical literature. It is important to integrate the findings from these distinct perspectives and devise more comprehensive intervention approaches that reflect the full state of knowledge so as to guide both clinical practice with working teens, as well as community advocacy. Although adolescents use health care less than people in other age groups, young people between 15 and 24 made >70 million estimated outpatient visits in 2004, with 79 visits for preventive care among every 100 persons. As a result, there is considerable opportunity for pediatricians to interact with them around work-related issues.

Physicians should also be familiar with the youth labor laws in their locales so they can help their adolescent patients and their parents make wise decisions about jobs. In addition, knowing about the working environments of these youth may help physicians gain insight into other adolescent health and behavioral con-
cerns, such as fatigue, school performance, and peer interactions. Special attention is warranted when youth are working long, late hours, without adequate adult supervision or with minimal training. Adding information about teen worker rights and responsibilities to standard patient care is one way that pediatricians can begin to inform both parents and teens about the potential hazards of youth work to both health and educational performance. Excellent information is available from the federal government, state agencies, and non-profit organizations. Several useful Web sites appear in Table 5.

In addition, individual clinicians and professional organizations should develop the skills and practice of engaging with other health professionals in addressing the policy issues surrounding young worker safety as they have successfully done in other aspects of child and adolescent health and safety, including regulations for fire safe sleepwear, poison prevention packaging, regulation of unsafe infant products such as infant walkers, child passenger safety, and graduated driver licensing for teens. Working conditions for teens can include many hazards, some of which are addressed by federal or state child labor policies or by OSHA regulations and some that are not. Even when policies govern safety, implementation is sometimes incomplete, suggesting the need for not only careful monitoring of the policies themselves but also of improvements in policy implementation and enforcement, as with drunk driving laws or policies governing weapon carrying at school or graduated driver’s licensing.

ACKNOWLEDGMENTS
This project was supported by grant 5-RO1-OH03530-02 from the National Institute of Occupational Safety and Health.

We appreciate data-analysis assistance provided by Jose Sandoval, MS, Shankar Viswanathan, MS, and Xiaoyan Shi, MS, and help in instrument development from Darlene Adkins, Thomas Harris, JD, Chris Miara, MS, Susan Gallagher, MPH, Marlee Gurrera, MS, Linda Treiber, PhD, and Susan Pollack, MD, MPH. Carrie Johnston and Kimberly Rauscher, ScD, assisted with preparation of the manuscript.

REFERENCES

---

**TABLE 5 Resources for Additional Information**

<table>
<thead>
<tr>
<th>Resource</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federal Network for Young Workers Safety and Health (<a href="http://www.cdc.gov/niosh/fednet">www.cdc.gov/niosh/fednet</a>)</td>
<td></td>
</tr>
<tr>
<td>US Department of Labor, OSHA Teen Workers Safety and Health Page (<a href="http://www.osha.gov/SLTC/teenworkers/index.html">www.osha.gov/SLTC/teenworkers/index.html</a>)</td>
<td></td>
</tr>
<tr>
<td>Occupational Health Surveillance Program, Massachusetts Department of Public Health Teens at Work Surveillance and Prevention Project (<a href="http://www.state.ma.us/dph/bhsre/ohsp">www.state.ma.us/dph/bhsre/ohsp</a>)</td>
<td></td>
</tr>
<tr>
<td>Labor Occupational Health Program, UC Berkeley, Young Workers’ Health and Safety (<a href="http://ist-socrates.berkeley.edu/~safejobs">http://ist-socrates.berkeley.edu/~safejobs</a>)</td>
<td></td>
</tr>
<tr>
<td>National Council for Occupational Safety and Health (<a href="http://www.cosnetwork.org/english_resources.htm">www.cosnetwork.org/english_resources.htm</a>)</td>
<td></td>
</tr>
<tr>
<td>National Consumer’s League: Child Labor Coalition (<a href="http://ncnet.org/childlabor">http://ncnet.org/childlabor</a>)</td>
<td></td>
</tr>
</tbody>
</table>
The ketogenic diet, a high fat, adequate protein, low carbohydrate diet, has, during the past decade, had a resurgence of interest for the treatment of difficult-to-control seizures in children. This review traces its history, reviews its uses and side effects, and discusses possible alternatives and the diet’s possible mechanisms of action. Finally, this review looks toward possible future uses of the ketogenic diet for conditions other than epilepsy.
The ketogenic diet (KD), developed in the early 1920s, had fallen into disuse during the 1970s and 1980s with the rapid development of new anticonvulsant agents for epilepsy. The recent resurgence of interest and use of the diet can be dated to the American Epilepsy Society’s meeting in 1996. Currently, it is perhaps more effective than most of the newer medications.

The rediscovery of this effective therapy for childhood epilepsy has, within the past decade, had a major impact on the most difficult-to-control seizures of childhood and promises to have an impact on adults with epilepsy as well. It will, perhaps, be used for other medical conditions as well. New research into its mechanism of action shows promise in changing our thinking about cerebral metabolism and our understanding of the control of epilepsy. This review was intended to document and summarize the remarkable progress in the use and understanding of the diet over the past 10 years. However, it is important to first understand its past.

**HISTORY OF THE KD**

**1920–1990**

In the early 1920s, epilepsy was treated with the bromides and phenobarbital. Both drugs had major sedating adverse effects and were frequently ineffective in completely controlling seizures. Hugh Conklin, an osteopathic physician and faith healer in Battle Creek, Michigan, believed, without any evidence, that epilepsy was attributable to intoxication of the brain from substances coming from the intestine. He postulated that putting the intestine completely at rest would rid the body of the intoxication, and he thereby developed his “fasting” or “water treatment” for epilepsy. This treatment deprived the children with epilepsy of all food, giving nothing but water for as long as 25 days. In 1922, he reported a high percentage of cures, and many more children were free of seizures for prolonged periods of time. But even before Conklin’s report, word of his successful treatment spread to others in more mainstream medicine.

These reports promised hope for children with epilepsy and set off a flurry of clinical and research activity. Studies of the metabolic changes during fasting were undertaken in an attempt to understand the interrelationships of fat, protein, and carbohydrate metabolism. A review article at that time stated that ketone bodies caused by starvation were the immediate result of the oxidation of certain acids in the absence of sufficient glucose and postulated that they were anticonvulsant.

Although prolonged fasting was difficult for those with severe epilepsy, it was far better than the constant seizures. The first article to suggest that a diet high in fat and low in carbohydrates could simulate the metabolic effects of starvation was published in 1921 from the Mayo Clinic. This diet provided adequate protein for growth, minimal carbohydrate, and the remainder of the calories as fat and was virtually identical to the KD that is used today.

Reports of the effectiveness of this new “ketogenic” diet appeared throughout the next 2 decades until phenytoin (Dilantin) was discovered in 1938, and the attention of physicians and epilepsy researchers turned from the mechanisms of action and efficacy of the diet to the development and mechanisms of action of new anticonvulsant agents. The era of medication treatment for epilepsy had begun, and the KD, then thought to be relatively difficult, rigid, and expensive, fell by the wayside. Encouraged by the drug companies, physicians believed that new and more effective medications were on the horizon. As pediatric neurologists and epileptologists had less and less experience with the “classical” KD, fewer children were started on it, and fewer dieticians were trained in its rigors and nuances. Therefore, the diets prescribed were often less precise, rigorous, and effective than in previous years. These failures led to the widespread opinion that the diet did not work and was very difficult to tolerate. A review of the KD in 1995 cited the feelings of many physicians that the KD was no longer justified.

**The Early 1990s**

The KD had continued to be implemented in ~10 children each year at Johns Hopkins Hospital, initially under the direction of Dr Samuel Livingston and, subsequently, Dr John Freeman and their dietician Millicent Kelley. In the late 1980s, in response to a challenge from a recently graduated nutritionist who asked if the diet was still effective in this “era of anticonvulsant medications,” Kinsman reviewed 58 recent patients and found that despite the use of many new anticonvulsant medications, patients with refractory seizures had the same success rate with the KD as had been reported decades earlier.

The start of the new era of the KD began with a Hollywood producer, Jim Abrahams, and his son Charlie, who was incapacitated by uncontrollable seizures that were refractory to multiple medications and other treatments. Reading about the diet, Abrahams brought his son to Johns Hopkins Hospital, and the child’s seizures were stopped completely soon after starting the diet. To make other parents aware of the KD, Abrahams created the Charlie Foundation, which published a book about the KD, now in its fourth edition, and made a film about the diet for parents and physicians. Abrahams filmed a network television program (Dateline) about the diet in 1994, and Abrahams created a made-for-television movie, “First Do No Harm,” starring Meryl Streep. In anticipation of these events, the Charlie Foundation funded a 7-center study of the diet designed to allow these centers to treat the patients resulting from the anticipated publicity. The multicenter study was started in 1994 and presented to the American Epilepsy...
Society in 1996. The reports from the multicenter study\textsuperscript{12} and of 150 patients from Johns Hopkins\textsuperscript{13,14} were the first of an avalanche of abstracts and articles on the clinical outcomes of children who were treated with the diet, including outcomes of various aspects and modifications of the diet. The amazing increase in the number of clinical abstracts presented at the American Epilepsy Society meetings is shown in Fig 1. Considering the time involved for new anticonvulsant medications to be developed, investigated, marketed, and then widely accepted, it is difficult to imagine now that just 10 years ago was the first time in recent years that an abstract on the KD was presented at the American Epilepsy Society.

1996–2006: The Explosion of Interest and Studies

Efficacy Demonstrated

Authors from Johns Hopkins reported the outcomes of 150 consecutive children 3, 6, and 12 months after initiating the diet,\textsuperscript{13} as well as their 3- to 6-year follow-up\textsuperscript{14} (Table 1). With an intention-to-treat methodology, these 150 children (who had averaged 410 seizures per month and whose seizures had failed to adequately improve on a mean of 6.2 medications) had a dramatic outcome. Twelve months after initiating the diet, 7% of the children were seizure free, and another 20% had a 90% decrease in seizures. Three to 6 years later, 27% of these same children had few or no seizures. Most of them were now off the diet and on fewer or even no medications.

Since the 1920s, reports of efficacy have been remarkably consistent across all age groups, seizure frequencies, and international locations.\textsuperscript{15–18} In general, 10%–15% of children who initiated the diet were seizure free 1 year later, 30% had a >90% reduction in seizures, and 40% to 50% found that the diet was either too difficult to continue or insufficiently effective and therefore discontinued it during the first 6 months.

Although over the past decade there has been a dramatic increase in the number of anticonvulsant drugs available, the KD continues to demonstrate a higher degree of effectiveness, even in children whose seizures are refractory to these newer medications. Several recent meta-analyses have examined the publications about the diet and, although finding a lack of prospective controlled studies, found an enormous amount of prospective uncontrolled and retrospective evidence.\textsuperscript{15–18} Nearly all reviews have stated that despite the lack of class I and II data, the scientific basis for the diet is strong and future studies should help identify ideal candidates and ways to improve tolerability rather than solely to prove the diet’s efficacy in a controlled manner.\textsuperscript{18}

Rise in Usage Internationally

Ten years ago the KD was nearly unknown internationally. In the past 8 years there has been a dramatic increase in its use worldwide, and currently ~75 centers in 45 countries offer the KD.\textsuperscript{19} With the exception of parts of Central America and Africa, parents only need look to their country or a neighbor for a KD center. There have been sponsored conferences and symposia in Canada, Croatia, Cuba, England, Germany, Greece, India, and Italy in the past 2 years alone.

Cultural, religious, and financial differences among these centers have led to differences in approaches to providing the KD. Some use less or no fasting, some use different ratios (to encompass more rice and less fat in some countries in the East), and some allow increased fluid and calorie consumption.\textsuperscript{19} These practices have led to insights into other methods for providing the KD that will be discussed later in this review. It should be noted that especially in developing nations, the KD may be a cost-effective epilepsy therapy when compared with the rising costs of anticonvulsant medications.

Challenges and Changes to the Traditional Diet Protocol

The Johns Hopkins protocol\textsuperscript{10} for initiating and maintaining the KD has been gradually modified both at Johns Hopkins and other centers and is continually evolving. Highlights of its evolution and changes are shown in Table 2.
Recent studies have demonstrated that the diet may not require a fasting phase at initiation and may be initiated either with full calories\(^\text{20,21}\) or without an admission to the hospital.\(^\text{22,23}\) However, many centers still use at least some fasting periods because of the occasional immediate benefits seen for some children. We have observed a dramatic effect of fasting similar to a “loading” dose of intravenous anticonvulsant agents.\(^\text{24}\) Most centers still admit children to observe their initial response to the diet and possible immediate adverse effects.

### TABLE 1
Outcomes of a 150-Patient Cohort Using the KD at Johns Hopkins Hospital

<table>
<thead>
<tr>
<th>Seizure reduction, n (%)</th>
<th>3 mo</th>
<th>6 mo</th>
<th>12 mo</th>
<th>3–6 y and Longer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seizure free</td>
<td>4 (3)</td>
<td>5 (3)</td>
<td>11 (7)</td>
<td>20 (13)</td>
</tr>
<tr>
<td>90%–99%</td>
<td>46 (31)</td>
<td>43 (29)</td>
<td>20 (20)</td>
<td>21 (14)</td>
</tr>
<tr>
<td>50%–89%</td>
<td>39 (26)</td>
<td>29 (19)</td>
<td>34 (23)</td>
<td>24 (16)</td>
</tr>
<tr>
<td>&lt;50%</td>
<td>36 (24)</td>
<td>29 (19)</td>
<td>8 (5)</td>
<td>18 (12)</td>
</tr>
<tr>
<td>No. (%)(% remaining on the KD</td>
<td>125 (83)</td>
<td>106 (71)</td>
<td>83 (55)</td>
<td>18 (12)</td>
</tr>
</tbody>
</table>


### TABLE 2
Factors Involved in KD Initiation and Maintenance, With Changes Described

<table>
<thead>
<tr>
<th>Diet Factor</th>
<th>Management</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diet initiation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950s</td>
<td>Admission and fasting until the child had lost 10% of body weight and had urinary ketones of 160 mg/dL</td>
<td></td>
</tr>
<tr>
<td>1960s–1990s</td>
<td>Admission and fasting 48 h</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>Johns Hopkins: admission and fasting 24 h ± ketosis, calories gradually increased; other centers have demonstrated efficacy without a fast and with full calories immediately; most centers still admit for education and immediate adverse-effect monitoring</td>
<td>Value of admission and fasting unproven; Atkins and low glycemic index treatment studies do neither</td>
</tr>
<tr>
<td><strong>Modified Atkins diet and low glycemic index under study, both 1:1 ratio diets</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>3:1 infants and adolescents</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>Generally still the same</td>
<td>Many countries in Asia use lower ratios</td>
</tr>
<tr>
<td>Atkins diet under study, probably 2:1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston: low glycemic index treatment even lower</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fluids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>Tight restriction to 80% daily requirement</td>
<td>Most children before the diet were drinking less fluids than either the daily requirement or the diet allotment</td>
</tr>
<tr>
<td>Current</td>
<td>Unclear if fluid restriction is necessary</td>
<td>May be helpful to prevent kidney stones by giving ad lib</td>
</tr>
<tr>
<td><strong>Calories</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>Tight restriction to 75% daily requirement, with modifications made afterward</td>
<td>Animal, not human, studies demonstrate benefit to caloric restriction</td>
</tr>
<tr>
<td>Current</td>
<td>Variable on the basis of estimated needs</td>
<td></td>
</tr>
<tr>
<td><strong>Carbohydrates</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>&lt;10 g/d</td>
<td>Atkins studies show that 20 g/d may be efficacious</td>
</tr>
<tr>
<td>Current</td>
<td>Unclear maximum amount</td>
<td></td>
</tr>
<tr>
<td><strong>Ketones</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>≧3 to ≧4 (80–160 mg/dL) urine ketones, checked frequently, believed crucial for seizure control</td>
<td>Interest in alternative methods of measurement (breath and serum)</td>
</tr>
<tr>
<td>Current</td>
<td>Necessary level and even importance unclear</td>
<td></td>
</tr>
<tr>
<td><strong>Duration of diet (maximum)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>2 y</td>
<td>Some evidence for patients with infantile spasms suggests that short periods (eg, 6 mo) may be sufficient</td>
</tr>
<tr>
<td>Current</td>
<td>As long as it is helpful</td>
<td></td>
</tr>
</tbody>
</table>
Alternative KDs

The medium-chain triglycerides (MCTs) diet uses fat sources that are more ketogenic than the saturated long-chain triglycerides (LCTs) typically consumed in the traditional KD, thus allowing more carbohydrates to be incorporated into the diet.25–27 A trial comparing the MCT diet, classic LCT diet, and a modification of the 2 (the Radcliff diet) found they were of roughly equal efficacy, but the MCT diet had a higher incidence of abdominal cramps, diarrhea, nausea, and vomiting.27 Occasionally we add small quantities of MCT oil to the classic KD to alleviate constipation and dyslipidemia. An ongoing trial is investigating the LCT and MCT diets in a randomized manner (J. H. Cross, MD, personal communication, 2006).

A modified Atkins diet also is emerging as a possible alternative dietary treatment for seizures.28,29 With restriction of carbohydrates (10–20 g per day), the Atkins diet can induce ketosis and does not restrict protein, fluid, or calories and does not require an admission or a fast. In a follow-up study, 65% of patients on the Atkins diet had a >50% reduction in seizures and 6 (35%) had a >90% reduction.29 Additional studies of the modified Atkins diet are underway including adult patients with epilepsy.

A third diet is the low glycemic index diet,30 in which fruits, breads, and starches are discouraged. This diet has even fewer carbohydrate restrictions than the modified Atkins diet.

Who Are the Best Candidates for the Diet?

The efficacy of the diet is independent of the type of seizure and is effective for both generalized and partial seizures31 at varied ages.31–35 Some refractory disorders that respond to the diet include Dravet syndrome,36 myoclonic-astatic epilepsy,37–39 Rett syndrome,40,41 migrational disorders,42 and tuberous sclerosis complex.43,44 The KD may be particularly helpful in the treatment of infantile spasms, especially when used earlier in the course of the disorder. Formula-based diets, whether fed via bottle or gastrostomy tube, have shown improved compliance as well as efficacy.46,47 Although no one particular anticonvulsant medication has been described as specifically beneficial in combination with the diet, the use of the diet in combination with vagus nerve stimulation may be synergistic.48

Patients with partial-onset (focal) seizures seem less likely to either improve significantly or become permanently seizure free.49,50 Children with a demonstrated, surgically approachable focus may respond to the diet but have less likelihood of either >90% seizure reduction or actual seizure freedom (unpublished data). In a limited case series, Lafora body disease did not respond well to the diet.51

Mechanisms of Action of the Diet

The mechanism(s) through which the KD exerts its anticonvulsant effects remains elusive. Although there is an abundance of data regarding the physiologic effects a KD exerts on humans and rodents, how these effects contribute to seizure protection is unclear. The diet has both anticonvulsant (ie, stopping a discrete seizure) and antiepileptic (ie, stopping the propensity to develop recurrent unprovoked seizures, or epilepsy) effects. The latter is suggested by series that examined the diet’s potential disease-modifying effects in patients who discontinued the diet, sometimes after only months, yet still enjoyed long-term freedom from seizures.14,52

Anticonvulsant effects of the KD have been studied primarily in models of nonepileptic rodents receiving a KD and later exposed to proconvulsant agents or electrical stimuli (eg, pentylentetrazol, maximal electroshock).53,54 Studies in mice that examined changes in glutamate (eg, the primary central nervous system excitatory neurotransmitter) and γ-aminobutyric acid (GABA) (eg, the primary central nervous system inhibitory neurotransmitter) suggested a key role for the KD in protection from seizures.55 Although actual levels of these neurotransmitters may not be elevated, there is a suggestion that flux through the GABA shunt may be increased, thus favoring inhibition of aberrant neuronal firing.56 Changes in levels of GABA (measured by magnetic resonance spectroscopy) and other cerebrospinal fluid amino acids have been documented in patients on the KD, which suggests that they may play a role in seizure protection.57,58

Changes in mitochondrial biogenesis (eg, increased metabolic enzymes and mitochondrial number) also have been documented, reviving an old hypothesis that changes in cellular metabolism may modify the cellular milieu into a less hyperexcitable (and hence, less epileptiform) state.54,59 Early work on the KD suggested that ketone bodies, especially β-hydroxybutyrate, might be anticonvulsant.25 Subsequent in vitro work argued against this idea, but acetone (another ketone body), initially believed to be unimportant because of its volatility, has anticonvulsant properties.60–62

The KD may also have antiepileptic effects. One speculation is that the antiepileptic effect is exerted via neuroprotection, but the mechanism for this is unclear. Neuroprotection may involve either protection from free oxygen radicals or prevention of apoptosis. Protection from free radicals may be provided via a decrease in coenzyme Q semiquinone,63 elevated mitochondrial uncoupling proteins, or elevated glutathione peroxidase. Uncoupling proteins, shown to be induced in mice that consume a KD, dissipate the mitochondrial membrane potential, thus protecting against free radical damage.64 The mechanism of this induction might be via fatty acids, which are elevated in the serum of patients on the
KD. The KD also induces glutathione peroxidase, which subsequently prevents damage to the cell membrane caused by lipid peroxidation. The KD may protect against apoptosis via increased levels of the protective protein calbindin or prevention of the accumulation of the pro–cell-death protein clusterin. Very recent work has shown the role of 2-deoxyglucose, an inhibitor of glycolysis, in protection from seizures. Reports of its anticonvulsant and antiepileptic properties suggest that there may be antiglycolytic compounds which may possibly mimic some of the mechanisms of action of the KD and constitute a new class of drugs for treating epilepsy. Additional investigations into the mechanism(s) of action of the KD may lead to answers not only as to why it works but also what may cause seizures and epilepsy to develop.

**Adverse Effects of the Diet**

Adverse effects of the KD only infrequently require the diet to be discontinued but are important for neurologists and pediatricians to recognize. Early-onset adverse effects associated with diet initiation include acidosis, hypoglycemia, gastrointestinal distress, dehydration, and lethargy. They are typically transient and easily managed and are minimized if patients are not fasted. Later adverse effects include dyslipidemia, kidney stones, and slowing of growth.

Cholesterol and lipids are adversely affected on the diet. The most extensive study of dyslipidemia on the diet followed 141 children prospectively over 2 years. In these children, there was an increase in atherogenic apoB-containing lipoproteins very low-density lipoprotein and low-density lipoprotein and a decrease in the antiatherogenic high-density lipoprotein cholesterol. Cholesterol increased ~130% but then stabilized over the 2-year period. It is interesting to note that the lipid profiles of children on the KD > 6 years returned toward baseline. The long-term effects of these changes in lipids, if any, are unknown, but it should be recalled that most patients remain on the diet for only 2 years and then return to a diet with normal fat ingestion.

Kidney stones occur in 5% of children on the KD and is thought to be secondary to a combination of acidosis, urine acidification, hypercalciuria, and hypocitraturia. Although anticonvulsant agents with carbonic anhydrase-inhibition properties (topiramate and zonisamide) have an independent risk of stones, the combined prevalence with the diet was not higher than either therapy alone. The risk of stones has significantly decreased since the prophylactic use of oral potassium citrate (Polycitra K) to alkalinate the urine. Children on the diet grow normally, but the growth of younger children seems to be slowed more than that of older children. Those on the diet for > 6 years were typically in the < 10th percentile for height and weight. Growth seems to increase rapidly after diet discontinuation. Children on the KD are monitored carefully and regularly by a registered dietitian for weight and height slowing.

Bone density may be decreased by the KD. A higher risk of skeletal fractures in children on the KD has been reported. Prevention with calcium supplementation, pamidronate, or lower KD ratios remain unproven.

Deaths have been reported in patients on the diet, although it is unclear that any of the deaths have been a result of the diet.

It is clear that the success and safety of the diet are best achieved by the close supervision of the patient by an experienced team that includes the physician, the dietician, and, often, a nurse.

**Other Uses Beyond Epilepsy**

Multiple uses of the KD are being investigated. All reported studies to date are very preliminary but are discussed in this review to indicate possible future uses of the diet. Neurodegenerative disorders provide a unique opportunity to study cellular protection via dietary means. In fact, the mechanism of the KD in neuroprotection might be more straightforward than its mechanism of protection against seizures.

Parkinson disease may be partly attributable to dysfunction of mitochondrial complex I. The KD, which effectively bypasses the requirement for complex I, may provide an alternative pathway for normal cellular metabolism. This notion served as the rationale for a case series that showed some improvement in clinical rating scales in 7 adults with Parkinson disease who consumed a KD for 28 days. Although the hypothesis that bypassing complex I is an attractive one, it is possible that lower dietary protein levels and weight loss in patients on the diet simply improved the patients’ baseline levodopa pharmacokinetics.

A KD reduced amyloid-β 40 and 42 in a mouse model of Alzheimer disease. Administration of a KD for 43 days was associated with a decrease in the amount of total brain amyloid-β content, although performance on an object-recognition task was unchanged. The KD also was associated with delayed progressive motor neuron loss and improved performance on a motor task (compared with controls) in a transgenic mouse model of amyotrophic lateral sclerosis, with in vitro data again showing a protective effect of β-hydroxybutyrate. The KD was associated with decreased cortical contusion volumes 7 days after a standardized controlled cortical impact in rats of specific pediatric ages. A key role for ketone bodies was suggested in a study that showed improved adenosine triphosphate production in the same trauma model after an infusion of β-hydroxybutyrate.

The KD has been reported to have decreased tumor size in 2 patients with astrocytomas and recently has been shown to inhibit brain tumor growth in a mouse model of an astrocytoma. Ketosis may also improve...
migraine headaches in a manner similar to the beneficial preventive effect of many anticonvulsant drugs.\textsuperscript{88} The Atkins diet has been reportedly effective for narcolepsy in a small case series, as well.\textsuperscript{89}

Psychiatric disorders have also been treated with the KD. Recent studies have demonstrated its use for autism and depression.\textsuperscript{90,91} The mechanism of action of the KD for psychiatric disorders is unclear.

In addition, the diet may have uses beyond neurologic disorders. Very preliminary studies indicate that the KD may be useful in conditions that involve an imbalance of glucose metabolism, including type 2 diabetes mellitus and polycystic ovary syndrome.\textsuperscript{92,93} The diet has been described for use in hypercholesterolemia.\textsuperscript{94}

CONCLUSIONS

The past decade has been an amazing one for those interested in the KD. Its increasing use in children with difficult-to-control seizures has opened new vistas for these children and also for our understanding of epilepsy. Its potential use in adults by using a less restrictive Atkins diet may make a difference to this population as well. The diet’s documented efficacy and tolerability have opened new horizons as it is tried for a variety of ills from brain tumors to migraine, and from head trauma to neurodegenerative diseases. Most exciting is the realization that beliefs concerning a high-fat diet making people fat and dyslipidemic have been proven false. Researchers are rediscovering that ketone bodies are not necessarily bad and that glucose is not necessarily good. A whole new era of metabolic research has opened up. It is not completely clear where it will lead, but its promise is exciting.

ACKNOWLEDGMENT

Dr Hartman was supported by the Epilepsy Foundation through the generous support of Pfizer, Inc.

REFERENCES

7. Whoese JW. The ketogenic diet: fa(c)it or fiction. \textit{J Child Neurol.} 1995;10:419–423
11. Abrahams J. \textit{An Introduction to the Ketogenic Diet: A Treatment for Pediatric Epilepsy} [videotape]. Santa Monica, CA: Charlie Foundation; 2000
73. Sampath A, Kossoff EH, Furth SL, Pyzik PL, Vining EP. Kidney
76. Liu TC, Megan P, Campbell K, Curtis R. A retrospective study: growth status of children with epilepsy post treatment with the ketogenic diet [abstract]. Epilepsia. 2005;46(suppl 8):155

MOST STUDENTS IN BIG CITIES LAG BADLY IN BASIC SCIENCE

“At least half of eighth graders tested in science failed to demonstrate even a basic understanding of it in 9 of 10 major cities, and fourth graders, the only other group tested, fared little better, according to results released here Wednesday. The outcome of those tests, part of the National Assessment of Educational Progress, often called the nation’s report card, showed that student performance in urban public schools was not only poor but also far short of science scores in the nation as a whole.”


Noted by JFL, MD
Has Blood Pressure Increased in Children in Response to the Obesity Epidemic?

Arnoud Chiolero, MD, MSc*, Pascal Bovet, MD, MPH*, Gilles Paradis, MD, MSc†, Fred Paccaud, MD, MSc*

*Community Prevention Unit, Institute of Social and Preventive Medicine, University of Lausanne, Lausanne, Switzerland; †Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

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

ABSTRACT

The associations between elevated blood pressure and overweight, on one hand, and the increasing prevalence over time of pediatric overweight, on the other hand, suggest that the prevalence of elevated blood pressure could have increased in children over the last few decades. In this article we review the epidemiologic evidence available on the prevalence of elevated blood pressure in children and trends over time. On the basis of the few large population-based surveys available, the prevalence of elevated blood pressure is fairly high in several populations, whereas there is little direct evidence that blood pressure has increased during the past few decades despite the concomitant epidemic of pediatric overweight. However, a definite conclusion cannot be drawn yet because of the paucity of epidemiologic studies that have assessed blood pressure trends in the same populations and the lack of standardized methods used for the measurement of blood pressure and the definition of elevated blood pressure in children. Additional studies should examine if favorable secular trends in other determinants of blood pressure (eg, dietary factors, birth weight, etc) may have attenuated the apparently limited impact of the epidemic of overweight on blood pressure in children.
Elevated blood pressure (BP) is one of the leading contributors to the global disease burden worldwide and accounts for 7 million deaths each year.1-2 Several causes of hypertension operate early in life.3-14 In particular, BP relates, in children as in adults, to body weight,6 dietary factors,7 physical activity,8 and birth weight.9

Until recently, hypertension was considered to be a rare condition in children.10 However, because the worldwide prevalence of pediatric overweight has increased largely over the last 2 decades11-13 and the association between body weight and BP is well documented in children,3,14-16 it is generally believed that BP in children should have increased in parallel to the overweight epidemic.6,17

Although the prevalence of type 2 diabetes among children has risen worldwide in parallel to the epidemic of overweight,18,19 few reliable population-based data are available to document trends in BP in children and adolescents. A review on trends from 1948 to 1998 reported a decline in BP in high-income countries in children, adolescents, and young adults aged 5 to 34 years.20 Whether the increasing prevalence of overweight in children has resulted in a commensurate increase in BP is an important issue. Elevated BP acquired in childhood tends to track into adulthood,3 and factors that affect BP in childhood are likely to further increase the burden of hypertension-related diseases in adults.

In this article, we aimed to (1) discuss methodologic and other issues that limit the validity of estimates of elevated BP in children, (2) review the prevalence of elevated BP in children and trends over time on the basis of recent methodologically sound population-based studies, and (3) examine if trends in elevated BP in children seem to reflect the rising prevalence of overweight.

METHODOLOGIC ISSUES

Limitations Related to BP Estimation in Children

There are several methodologic limitations that complicate the valid assessment of BP in children.

A first issue relates to the number of readings on which estimates of “elevated BP” are based. BP decreases over subsequent readings during a single visit as well as over readings taken on separate visits. This decrease partly reflects an alert reaction and is observed in both adults21 and children.15,16 It follows that BP obtained at one occasion tends to overestimate usual BP. Most epidemiologic studies have relied on only 1 set of BP readings obtained at a single visit. It is notable that the normative US reference BP data are based on the first BP reading obtained at 1 visit.22 However, in most recent surveys (Table 1), 2 or 3 BP readings were obtained and BP was based on either the mean of 2 readings,23,24 the mean of 3 readings,25 or the mean of the last 2 of 3 readings.14 Normative data recently proposed for English children were determined on the basis of the mean of the last 2 of 3 readings obtained during a single visit.26 However, when BP was measured on different visits, the prevalence of elevated BP decreased from 19.4% (first visit), to 9.5% (second visit), and to 4.5% (third visit) in 1 American study27 and from 8.8% (first visit) to 4.2% (second visit) in 1 European study.16

The use of different BP-measurement devices also limits the comparability of estimates of elevated BP across studies. Although mercury manometers were used in the past, automated oscillometric devices are now used increasingly often. Electronic devices have several advantages including a greater reliability and the avoidance of observer bias.27 However, proper validation and calibration of such devices are necessary because substantial systematic bias can exist between different models and manufacturers.28-30 Unfortunately, few devices have been validated for use in children.31 Moreover, current validation protocols are quite lenient and allow, for example, for systematic underestimates or overestimates of BP by as much as 5 mm Hg.32

The use of the auscultatory technique (mercury or aneroid) raises specific issues.33 Rounding errors, bias related to expected values (which may be stronger in children), and other operator-dependant biases can occur.34 The accuracy of aneroid sphygmomanometers needs to be checked regularly. Moreover, diastolic BP may be difficult to determine when the Korotkoff sounds do not disappear, which can occur in younger children. The choice of either the fourth Korotkoff phase (K4: muffling of sound, eg, to overcome the problem of nondisappearance of Korotkoff sounds in some children) or the fifth phase (K5: disappearance of sound, as usually applied with adults) to define diastolic BP in children has important consequences on BP estimates.34 For example, in a study of school-aged girls, the mean difference in BP between K4 and K5 was 9.9 mm Hg.35 The US reference data are based on K5 to define diastolic BP.22

In addition to biases related to BP-measurement procedures, age-adjusted BP in children may be affected by secular trends in children’s height.36 BP relates strongly to height independently of age and gender22,37,38 so that secular increases in children’s height are likely to translate into some increase in mean BP levels over time. This issue has relevance in view of the largely different mean heights across populations and the substantial increase in children’s height over time in many populations. In several recent studies of BP trends in the United States25,39 and United Kingdom,40 the authors did not adjust BP for height, and part of the observed increase in BP over time might relate to a concomitant increase in children’s height. In an American study, an observed increase in BP between 1986 and 1996 was largely reduced when adjusted for weight and height.41 The potential confounding effect of height may be especially strong in developing countries, where age-specific
# Prevalence of Elevated BP in Recent Large Population-Based Surveys of Children and Adolescents

<table>
<thead>
<tr>
<th>Study (Year) and Sample</th>
<th>Age, y</th>
<th>Year</th>
<th>N</th>
<th>Mean BMI or Prevalence of Overweight&lt;sup&gt;a,b&lt;/sup&gt;</th>
<th>Device</th>
<th>Set of BP Readings</th>
<th>Definition of BP</th>
<th>Proportion With Elevated Systolic BP, %&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Proportion With Elevated Diastolic BP, %&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Proportion With Elevated BP, %&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adrogué and Sinaiko&lt;sup&gt;23&lt;/sup&gt; (2001); schoolchildren of cities of St Paul and Minneapolis, MN</td>
<td>10–15</td>
<td>1986–1987</td>
<td>19 542</td>
<td>Boys: 20.5 kg/m&lt;sup&gt;2&lt;/sup&gt; Girls: 19.9 kg/m&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Mercury sphygmomanometer</td>
<td>First set</td>
<td>Mean of 2 readings</td>
<td>2.7</td>
<td>20</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Second set, if first BP ≥ 70th percentile, up to 21 d after first set</td>
<td>Mean of 2 readings</td>
<td>0.8</td>
<td>0.4</td>
<td>NA</td>
</tr>
<tr>
<td>Jafar et al&lt;sup&gt;24&lt;/sup&gt; (2005); national random sample, Pakistan</td>
<td>5–14</td>
<td>1990–1994</td>
<td>5641</td>
<td>Boys: 15.2 kg/m&lt;sup&gt;2&lt;/sup&gt; Girls: 15.3 kg/m&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Mercury sphygmomanometer</td>
<td>First set</td>
<td>Mean of 2 readings</td>
<td>8.7</td>
<td>96</td>
<td>15.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Second set</td>
<td>Mean of 2 readings</td>
<td>6.2</td>
<td>32</td>
<td>8.7</td>
</tr>
<tr>
<td>Paradis et al&lt;sup&gt;25&lt;/sup&gt; (2004); random sample, Province of Quebec, Canada</td>
<td>9, 13, and 16</td>
<td>1999</td>
<td>3589</td>
<td>Boys and girls, 9 y: 9%&lt;sup&gt;a&lt;/sup&gt; Boys and girls, 13 y: 9%&lt;sup&gt;a&lt;/sup&gt; Boys and girls, 16 y: 8%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Automated oscillometric Dynamap XL CR 9340</td>
<td>First set</td>
<td>Mean of last 2 of 3 readings</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Second set</td>
<td>Mean of 3 readings</td>
<td>13</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Third set</td>
<td>Mean of 3 readings</td>
<td>17</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Sorof et al&lt;sup&gt;26&lt;/sup&gt; (2002); children of 8 schools in Houston, TX</td>
<td>12–16</td>
<td>2000–2001</td>
<td>2460</td>
<td>Boys and girls: 23%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Automated oscillometric SpaceLabs monitor</td>
<td>First set</td>
<td>Mean of 3 readings</td>
<td>16</td>
<td>2</td>
<td>16.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Second set, if first BP ≥ 90th percentile, 1–2 wk after first set</td>
<td>Mean of 3 readings</td>
<td>11</td>
<td>1</td>
<td>11.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Third set, if second BP ≥ 95th percentile, 1–2 wk after second set</td>
<td>Mean of 3 readings</td>
<td>8</td>
<td>2</td>
<td>9.5</td>
</tr>
<tr>
<td>Sorof et al&lt;sup&gt;27&lt;/sup&gt; (2004); children of 8 schools in Houston, TX</td>
<td>13.5 ± 1.7</td>
<td>2002</td>
<td>5102</td>
<td>Boys and girls: 20%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Automated oscillometric SpaceLabs monitor</td>
<td>First set</td>
<td>Mean of 3 readings</td>
<td>NA</td>
<td>NA</td>
<td>19.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Second set, if first BP ≥ 95th percentile, 8–14 d after first set</td>
<td>Mean of 3 readings</td>
<td>NA</td>
<td>NA</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Third set, if second BP ≥ 95th percentile, 15–21 d after first set</td>
<td>Mean of 3 readings</td>
<td>NA</td>
<td>NA</td>
<td>4.5</td>
</tr>
<tr>
<td>Genovesi et al&lt;sup&gt;28&lt;/sup&gt; (2005); children of the schools of 5 villages, Province of Milan, Italy</td>
<td>6–11</td>
<td>2003–2004</td>
<td>24 16</td>
<td>Boys: 24.7%&lt;sup&gt;b&lt;/sup&gt; Girls: 29.3%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Mercury sphygmomanometer</td>
<td>First set</td>
<td>1 reading</td>
<td>NA</td>
<td>NA</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Second set, if first BP ≥ 95th percentile, 8–15 d after first set</td>
<td>Mean of 3 readings</td>
<td>NA</td>
<td>NA</td>
<td>4.2</td>
</tr>
<tr>
<td>Falkner et al&lt;sup&gt;29&lt;/sup&gt; (2006); children visiting a pediatrician network, Delaware</td>
<td>2–19</td>
<td>2002</td>
<td>18 618</td>
<td>Boys and girls: 20.2%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Aneroid sphygmomanometer</td>
<td>First set</td>
<td>1 reading&lt;sup&gt;e&lt;/sup&gt;</td>
<td>6</td>
<td>2</td>
<td>7.2</td>
</tr>
<tr>
<td>Chiolero et al&lt;sup&gt;30&lt;/sup&gt; (2006); all schoolchildren of 4 grades, Republic of Seychelles</td>
<td>5–16</td>
<td>2002–2004</td>
<td>15 612</td>
<td>Boys: 7.5%&lt;sup&gt;a&lt;/sup&gt; Girls: 7.8%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Automated oscillometric Omron M5</td>
<td>First set</td>
<td>Mean of 2 readings</td>
<td>6.1</td>
<td>5.1</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Second set</td>
<td>Mean of 2 readings</td>
<td>6.0</td>
<td>6.9</td>
<td>10.1</td>
</tr>
</tbody>
</table>

All studies were school based except for the study by Falkner et al. <sup>29</sup> NA indicates not available.

Overweight refers to either <sup>a</sup> BMI ≥ 95th gender-, age-, and height-specific US percentiles (determined in the 1970s and 1980s) or <sup>b</sup> the criteria for overweight of the International Obesity Task Force.

<sup>c</sup> Elevated BP: BP ≥ 95th gender-, age-, and height-specific US percentiles (determined in the 1970s and 1980s).<sup>22</sup>
<sup>d</sup> Percentiles were based on the current BP distribution of the participants.

<sup>e</sup> Data are from medical charts; most often, 1 BP reading was taken.
height of children may increase rapidly with socioeconomic development.

Current US BP reference data are determined on gender-, age-, and height-adjusted percentiles to take body size into account.2 Reference values have also been proposed for ambulatory BP and are provided according to body size directly (by height increments of 5 cm).42

**Definition of Elevated BP in Children**

Observational studies and clinical trials in adults indicate that BP is associated with both cardiovascular diseases (CVDs) and total mortality43 and that BP reduction lowers CVD risk.44 Above 115/75 mm Hg, the risk of CVD doubles with each increment of 20 mm Hg systolic or 10 mm Hg diastolic BP.45 However, because the relationship between BP and CVD risk is graded over the entire range of BP, a definition of hypertension is inherently arbitrary. Rose proposed an operational definition of hypertension: “the level [of BP] at which the benefits (. . .) of action exceed those of inaction.”46 Among adults, hypertension is currently defined as a sustained BP ≥140/90 mm Hg and/or current use of antihypertensive treatment.45

In children, no cohort data are available to relate BP with CVD mortality or morbidity.17,47 However, elevated BP in children is associated with several intermediate outcomes such as ventricular hypertrophy48 or increased carotid intima-media thickness.49–52 Furthermore, BP tracks from childhood into adulthood, which suggests that elevated BP occurring as early as in childhood may worsen CVD risk in adults. In the Bogalusa Heart Study, BP at age 5 to 14 correlated with BP at age 20 to 31 (correlation coefficients of 0.36–0.50 for systolic BP and 0.20–0.42 for diastolic BP).53

The most widely used cutoff values for defining elevated BP in children are based on BP percentiles specific for gender, age (1-year intervals), and height (7 categories based on height percentiles).2 These norms were determined on the basis of data in 63,227 American children aged 1 to 17 years who were participating in a variety of studies or surveys in the 1970s and 1980s (ie, before the current obesity epidemic)12 and were based on the first BP reading taken during a single occasion. “Elevated BP” is defined for BP values ≥95th gender-, age-, and height-specific percentile, and a child is considered to be “hypertensive” if he or she has elevated BP on at least 3 separate occasions.22

Pooled data from 6 Northwest European studies conducted in 1975–1986, which were based on the first reading taken during a single visit, showed that age-, gender-, and height-adjusted percentiles of systolic/diastolic BP were, on average, 6/3 mm Hg higher in European than in American children,64 which underscores the potential limit of using American reference data for European children.

**ELEVATED BP IN CHILDREN AND TRENDS OVER TIME**

A literature search was undertaken systematically by one of us (Dr Chiolero) by using the Medline/PubMed database with the search terms “children,” “blood pressure,” and “survey.” Chosen articles were restricted to those written in English and published between 1980 and February 2006. References of the selected articles were examined also.

First, we report below (in “Current Prevalence of Elevated BP in Children”) the prevalence of elevated BP that was estimated in the available population-based studies. We selected these population studies if (1) the prevalence of elevated BP was specifically addressed, (2) details about BP measurement were provided, and (3) US reference data were used22,38 to define BP percentiles and elevated BP (Table 1). In most of these surveys, elevated BP was defined as BP equal to or above the American gender-, age-, and height-specific 95th percentile (ie, calculated in the 1970s and 1980s, as explained above).22 In addition, the review was limited to large studies (≥2000 children). Large surveys have the power to provide prevalence estimates with some precision; because BP varies by age, gender, and height, prevalence of elevated BP must be examined across a large number of categories of age, gender, and height. Although this was not a requirement of the search strategy, most of these studies also defined overweight as a BMI equal to or above the American gender- and age-specific 95th percentile. It was noticeable that these reference BMI percentiles were determined on the basis of data gathered mostly between 1963 and 1980,55 a time period during which overweight was less frequent.12

Second, we report in “Recent Trends of BP in Children” trends on elevated BP based on single studies. For this purpose, studies were selected if they were based on a cohort study design or on paired surveys, which are suitable to indicate trends over time in mean BP levels or prevalence of elevated BP within defined populations and using comparable BP-measurement procedures (Table 2).

**Current Prevalence of Elevated BP in Children**

A convenience sample of American school-aged children aged 10 to 15 years were examined in 1986–1987 (Table 1).23,56 BP was measured with a mercury sphygmomanometer. At the first visit, the prevalence of elevated BP (≥95th percentile) was 2.0% to 2.7%. Students with BP >70th percentile were seen again on a separate day, and the prevalence of elevated BP decreased to 0.4% to 0.8%.

On the basis of BP measured during 1 visit in a representative sample, Pakistani children aged 5 to 14 years seen in 1990–1994 had higher BP than US children of the same age range who participated in the Third National Health and Nutrition Examination survey in 1988–1994 (NHANES III). This difference was found
Despite a mean BMI that was markedly (~3 kg/m²) lower in Pakistani than American children. The authors speculated that these differences could relate, at least partially, to lower birth weight in Pakistani than US children. Low birth weight is a known risk factor for subsequent elevated BP in children and in adults.

In Quebec, Canada, BP was measured 3 times during 1 visit in 1999 in a representative sample of children aged 9, 13, and 16 years. The prevalence of elevated BP (based on the American reference data) increased across age categories from 7% to 13% and 17%, respectively. The prevalence of overweight was 8% to 9%. Mean BMI was 4 to 6 kg/m² higher in children with systolic BP ≥95th percentile than in children with systolic BP <25th percentile. It was noticeable that most children with elevated BP had only elevated systolic BP. Less than 1% of children with elevated BP had elevated diastolic BP. Compared with data from an earlier Canadian survey conducted in 1978–1979, mean systolic BP was 4 to 8 mm Hg higher, whereas mean diastolic BP was 3 to 10 mm Hg lower. However, comparisons are limited because a mercury sphygmomanometer was used in the earlier survey, whereas an automated oscillometric device was used in the latter, and the 2 surveys also had different sampling designs.

### TABLE 2
Trends in BP in Children and Adolescents

<table>
<thead>
<tr>
<th>Study (Year) and Sample</th>
<th>Age, y</th>
<th>Gender</th>
<th>Year</th>
<th>N</th>
<th>Mean BMI or Prevalence of Overweight/Obesity b</th>
<th>Mean BP, mm Hg</th>
<th>Mean BP Difference, mm Hg</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gidding et al (1995); biracial samples, Bogalusa, LA</td>
<td>7–9</td>
<td>Boys</td>
<td>1973</td>
<td>109 m, 98 f</td>
<td>12.1 m, 12.2 f kg/m²</td>
<td>97/59 m, 96/60 f</td>
<td>0/0 m, 0/0 f</td>
<td>No adjustment</td>
</tr>
<tr>
<td></td>
<td>1984</td>
<td>49 m, 52 f</td>
<td>12.9 m, 12.6 f kg/m²</td>
<td>97/59 m, 96/60 f</td>
<td>0/0 m, 0/0 f</td>
<td>No adjustment</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1984</td>
<td>Girls</td>
<td>91 m, 119 f</td>
<td>12.8 m, 12.6 f kg/m²</td>
<td>94/57 m, 96/56 f</td>
<td>0/0 m, 0/0 f</td>
<td>No adjustment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15–17</td>
<td>Boys</td>
<td>1981</td>
<td>109 m, 98 f</td>
<td>12.5 m, 12.6 f kg/m²</td>
<td>114/69 m, 114/70 f</td>
<td>+1/2 m, +1/1 f</td>
<td>Rohrer index to assess adiposity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Girls</td>
<td>1981</td>
<td>91 m, 119 f</td>
<td>12.9 m, 13.5 f kg/m²</td>
<td>111/72 m, 112/72 f</td>
<td>−4/−1 m, +2/2 f</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1992</td>
<td>48 m, 52 f</td>
<td>13.5 m, 14.1 f kg/m²</td>
<td>110/68 m, 116/72 f</td>
<td>−4/−1 m, +2/2 f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Girls</td>
<td>1992</td>
<td>73 m, 62 f</td>
<td>14.2 m, 15.9 f kg/m²</td>
<td>105/70 m, 108/71 f</td>
<td>−5/−2 f, −4/−1 f</td>
<td></td>
</tr>
<tr>
<td>Muntner et al (2004); national random samples of schoolchildren, Princeton, NJ</td>
<td>10–14</td>
<td>Boys/girls</td>
<td>1973/1974</td>
<td>300</td>
<td>16.9 kg/m²</td>
<td>98/857.7</td>
<td>+2.9/+1.9</td>
<td>No adjustment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1989/1990</td>
<td>1286</td>
<td>18.3 kg/m²</td>
<td>101.7/59.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luepker et al (1999); schoolchildren, Minneapolis, MN</td>
<td>10–14</td>
<td>Boys</td>
<td>1986</td>
<td>4239</td>
<td>19.7 kg/m²</td>
<td>106.3/63.3</td>
<td>Adjusted for age, ethnicity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1996</td>
<td>5223</td>
<td>20.5 kg/m²</td>
<td>107.8/61.8</td>
<td>+1.5/−0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1998</td>
<td>3983</td>
<td>20.3 kg/m²</td>
<td>105.6/66.3</td>
<td>No adjustment for height</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1996</td>
<td>5018</td>
<td>21.2 kg/m²</td>
<td>106.3/64.2</td>
<td>+0.7/+1.9</td>
<td></td>
</tr>
<tr>
<td>Muntri et al (2004); English Health Survey; national random samples, US</td>
<td>8–17</td>
<td>Boys/girls</td>
<td>1988/1994</td>
<td>3496</td>
<td>11.7 kg/m²</td>
<td>104/6.584</td>
<td>Adjusted for age, gender, ethnicity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2000</td>
<td>2086</td>
<td>16.3 kg/m²</td>
<td>106.0/61.7</td>
<td>+1.4/+3.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2001</td>
<td>482</td>
<td>22.0 kg/m²</td>
<td>109.9/64.5</td>
<td>No adjustment for height</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2000</td>
<td>530</td>
<td>19.4 kg/m²</td>
<td>102.9/59.1</td>
<td>No adjustment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2001</td>
<td>514</td>
<td>20.3 kg/m²</td>
<td>104.2/60.4</td>
<td>No adjustment for height</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2000</td>
<td>2086</td>
<td>16.3 kg/m²</td>
<td>106.0/61.7</td>
<td>+1.4/+3.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2001</td>
<td>482</td>
<td>22.0 kg/m²</td>
<td>109.9/64.5</td>
<td>No adjustment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2000</td>
<td>2086</td>
<td>16.3 kg/m²</td>
<td>106.0/61.7</td>
<td>+1.4/+3.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2001</td>
<td>482</td>
<td>22.0 kg/m²</td>
<td>109.9/64.5</td>
<td>No adjustment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2000</td>
<td>2086</td>
<td>16.3 kg/m²</td>
<td>106.0/61.7</td>
<td>+1.4/+3.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2001</td>
<td>482</td>
<td>22.0 kg/m²</td>
<td>109.9/64.5</td>
<td>No adjustment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2000</td>
<td>2086</td>
<td>16.3 kg/m²</td>
<td>106.0/61.7</td>
<td>+1.4/+3.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2001</td>
<td>482</td>
<td>22.0 kg/m²</td>
<td>109.9/64.5</td>
<td>No adjustment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2000</td>
<td>2086</td>
<td>16.3 kg/m²</td>
<td>106.0/61.7</td>
<td>+1.4/+3.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1999/2001</td>
<td>482</td>
<td>22.0 kg/m²</td>
<td>109.9/64.5</td>
<td>No adjustment</td>
<td></td>
</tr>
</tbody>
</table>

**Note:**
- w indicates white; bl, black.
- Obesity refers to gender- and age-specific BMI criteria of the International Obesity Task Force.
- a Indicates white; b, black.
- b Obesity refers to gender- and age-specific BMI criteria of the International Obesity Task Force.

---

Despite a mean BMI that was markedly (~3 kg/m²) lower in Pakistani than American children. The authors speculated that these differences could relate, at least partially, to lower birth weight in Pakistani than US children. Low birth weight is a known risk factor for subsequent elevated BP in children and in adults.

In Quebec, Canada, BP was measured 3 times during 1 visit in 1999 in a representative sample of children aged 9, 13, and 16 years. The prevalence of elevated BP (based on the American reference data) increased across age categories from 7% to 13% and 17%, respectively. The prevalence of overweight was 8% to 9%. Mean BMI was 4 to 6 kg/m² higher in children with systolic BP ≥95th percentile than in children with systolic BP <25th percentile. It was noticeable that most children with elevated BP had only elevated systolic BP. Less than 1% of children with elevated BP had elevated diastolic BP. Compared with data from an earlier Canadian survey conducted in 1978–1979, mean systolic BP was 4 to 8 mm Hg higher, whereas mean diastolic BP was 3 to 10 mm Hg lower. However, comparisons are limited because a mercury sphygmomanometer was used in the earlier survey, whereas an automated oscillometric device was used in the latter, and the 2 surveys also had different sampling designs.

Two surveys were conducted at schools in Houston, Texas, in 2000/2001 and 2002. In both surveys, BP was measured 3 times during each of up to 3 separate visits if children had BP ≥95th percentile. In the first survey, the prevalence of overweight was 23% (BMI ≥95th percentile). The prevalence of elevated BP was 16.8% on the basis of readings from the first visit and 9.5% on the basis of readings from the third visit. As in the Canadian study mentioned above, the prevalence of...
elevated diastolic BP was very low. In the second Houston survey, the prevalence of overweight was 20%. The prevalence of elevated BP was 19.4% on the basis of the readings from the first visit and 4.5% on the basis of the readings from the third visit. BMI was strongly associated with systolic BP but not with diastolic BP. The prevalence of hypertension (defined as elevated BP at all 3 visits) was 10.7% in obese children (BMI ≥95th percentile) compared with 2.6% in children with normal BMI (<85th percentile).

In a school-based study performed in a representative sample of children aged 6 to 11 in 2004 in villages around Milan, Italy, the subjects with elevated BP at the first visit had BP measured again 8 to 15 days later. The prevalence of elevated BP decreased from 8.8% at the first visit to 4.2% at the second visit. Both systolic and diastolic BP were associated with BMI.

In Delaware, on the basis of electronic medical charts from children aged 2 to 19 attending primary care practices,60 the prevalence of elevated BP (based on only 1 BP reading for each subject) was surprisingly low (7.2%) despite a very large prevalence of overweight (20.2%; BMI ≥95th percentile). In this study, systolic and diastolic BP were related to BMI at all ages and among children aged <5 in particular.

As part of a school-based surveillance system in children of the Republic of Seychelles, a rapidly developing small island state in the African region with a large majority of the population of African descent, BP was measured in all children of 4 grades between ages 5 and 16 in 2002–2004.61,62 Systolic and diastolic BP were associated with BMI independent of gender, age, and height. In categories of children with normal weight (BMI <85th percentile), “at risk of overweight” (BMI ≥85th to <95th), and “overweight” (BMI ≥95th), proportions with elevated BP were 8%, 14%, and 23% in boys and 8%, 16%, and 29% in girls, respectively.

Overall, most of these recent studies reported a high prevalence of elevated BP. However, the prevalence seemed to depend substantially on the methodology used for BP measurement.

Recent Trends of BP in Children
In 2002, McCarron et al20 reviewed data on BP trends among individuals aged 5 to 34 years of both genders in developed countries. They concluded that BP declined between 1948 and 1998. For individuals aged 18 years or less, conclusions were based mostly on the NHANES for the period of 1963/80 and the English Health Surveys for the short period of 1995/1998.40 In the NHANES, BP was measured in American youth aged 6 to 17 years and was lower in 1971/1974 and 1976/1980 compared with 1963/1965. However, BP measurements differed between surveys in these periods: BP was measured in either the supine or sitting positions, the observer was either a physician or a nurse, and the number of readings differed between the surveys. In the NHANES 1971/1974 and 1976/1980 (all measuring with mercury devices), 40% to 50% of the readings had 0 end digits (vs 20% expected if BP was assessed along recommended 2-mm Hg increments).63

In the Bogalusa Heart Study, 2 convenience samples of white and black children aged 7 to 9 were examined in 1973 and 1984 (Table 2). The same children were examined again 8 years later in 1981 and 1992 at the age of 15 to 17. The mean of 6 BP readings was used, and the fourth Korotkoff sound was recorded for diastolic BP. Only small BP differences were observed among children aged 7 to 9 in 1973 and 1984. However, children aged 15 to 17 had lower systolic BP in 1992 compared with children of the same age examined in 1981 (except for black boys, who tended to have slightly higher BP in 1992). Only small differences were observed for diastolic BP over time.

In 1973/1974 and 1989/1990, BP was measured in convenience samples of children aged 10 to 14 in Princeton, Ohio.39 In 1973/1974, the reported BP was based on the average of 2 readings, whereas in 1989/1990 it was based on the average of the last 2 of 3 readings. Mean BP was higher in the latter than the former study, whereas the prevalence of elevated BP, surprisingly, tended to decrease. This result may reflect that the distribution of BP did not change uniformly: an increase in the lower range of BP could have occurred without concurrent change in the higher range of BP.

More recently, BP was measured by using the same methods in convenience samples of school-aged children aged 10 to 14 years in 1986 and in 1996 in Minneapolis, Minnesota.41 During the interval, systolic BP increased, but diastolic BP decreased. Adjustment for BMI largely eliminated the systolic BP increase between the 2 periods but did not alter the diastolic BP decrease over time.

In the recent NHANES surveys (1988/1994 and 1999/2000), BP was measured with similar procedures, and 3 BP readings were obtained from most children and adolescents aged 8 to 17 years33 (Table 2). After adjustment for age, race/ethnicity, and gender (but not for height), BP was slightly higher in 1999/2000 compared with 1988/1994. Most of the increase in BP over time could not be accounted for by concurrent trends in BMI: adjustment for BMI could explain only 29% of the increase in systolic BP and 12% of the increase in diastolic BP.

In the English Health Surveys,40 BP was measured by nurses at the children’s homes with an oscillometric automated device (Dinamap 8100; GE Healthcare, Critikon, FL). BP tended to decrease between 1995 and 1998 but increased slightly between 1998 and 2001. In 2004, BP was measured with a new automated device (Omron HEM 807; Omron Healthcare Europe BV, Hoofddorp, Netherlands), which is known to provide higher BP estimates compared with the Dinamap 8100. After data were calibrated to take this bias into account,
BP was only slightly higher in 2004 than in 2001 despite a continued increase in the prevalence of overweight.

In 1989/1990 and 1999/2001, BP was measured in a representative sample of Irish adolescents aged 12 and 15 years with a random-zero Hawksley sphygmomanometer. The mean of 2 BP readings was used in the first survey (1989/1990), but only 1 BP reading was measured in the second survey (1999/2001). Mean BMI increased significantly in children aged 12 but increased only marginally in children aged 15. In both age groups, BP decreased substantially between the 2 surveys. Adjustment for BMI, age, height, birth weight, and social class changed these estimates only marginally. Such a large decrease in BP is difficult to interpret. BP at the second survey was surprisingly low, particularly if one considers that it was based on 1 BP reading, which is usually a source of overestimation.

DISCUSSION

Most recent large population-based surveys have reported a relatively high prevalence of elevated BP in children and adolescents in several populations. However, the majority of the few available trend studies suggest that BP increased, at most, only moderately over time since the 1980s, which is in contrast to substantial increases in the prevalence of overweight reported in these same surveys. Some studies even showed a decrease in BP over time, both small (eg, in the Bogalusa Heart Study) and substantial (eg, in an Irish study). In addition, trends in systolic and diastolic BP were not always consistent. More generally, differences in the methods used to measure BP between studies (hence differences in the definition of “elevated BP”) strongly limited our ability to draw definitive conclusions on BP trends in children over the past few decades.

BMI (or any other index of adiposity) is a major determinant of BP, and the absence of a large increase in BP over time in most studies, despite a rising prevalence of overweight, suggests that other factors may also have influenced trends in BP over time. Such factors may include nutrition characteristics, such as the intake of fruits, vegetables, or dairy products or salt. These dietary factors have not been assessed in most of these studies, and so their independent effect on BP trends cannot be assessed. More generally, whereas the total caloric intake has increased largely in most countries worldwide, mixed trends (favorable and unfavorable) have been observed for specific nutrients that relate to hypertension, at least in US children.

Secular changes in nondietary factors may also have impacted on the trends in elevated BP over time. Low birth weight is associated with elevated BP in adults and children. Mean birth weight has increased throughout the last quarter century in many countries including the United States, Canada, and the United Kingdom and also in developing countries such as India despite a concomitant increase in preterm births in several of these countries. Increasing birth weight over time has been related to decreasing tobacco use among mothers during pregnancy, increased stature of mothers, favorable changes in socioeconomic factors, and other factors such as improved maternal nutrition. Increasing birth weight over the last decades may have accounted for some of the downward trends in BP over time, but no data currently support this possibility. Breastfeeding has been related to lower BP in children, which can be a result of the low salt content and high long-chain polyunsaturated fatty acid content of breast milk. However, during the 1990s, breastfeeding practices generally did not change substantially in many countries worldwide. Finally, in many countries, the amount of reported physical activity (particularly walking time or leisure exercise) has generally decreased. In addition to being a risk factor for obesity (itself related to elevated BP), low physical activity is associated independently with higher BP in children. However, precise measurement of all daily physical activity remains a challenge, as is reliable assessment of trends in physical activity over time.

Our findings are consistent with the review by McCarron et al, which reported a decline in BP from 1948 to 1998 in children, adolescents, and young adults in high-income countries. We did not perform a meta-analysis to estimate an average BP change over time in children because of excessively high heterogeneity of data in the available surveys and studies. First, studies differ widely in their particulars and design (ethnicity, sample, time period, and, most importantly, methodology for BP measurement and definition of elevated BP). Second, trends in BP over time may not be assumed to be similar in different populations in view of the different pace in the pediatric obesity epidemic and possibly also in differences in the relationship between BMI and BP between different populations.

CONCLUSIONS

The available data do not support the hypothesis that the worldwide epidemic of overweight in children has resulted in a commensurate increase in BP levels in children. However, none of the available studies on trends in BP over time has accounted for several other BP-related factors. Therefore, it is not possible to assess whether the apparent nonsystematic convergence of trends in elevated BP and overweight over time is related to a weakening in the relationship between BMI and BP over time or to concurrent secular changes over time in other risk/preventive factors of elevated BP. Furthermore, important differences in the methods used to assess BP in different studies limited our ability to directly compare results between populations and over time. This stresses the need for additional epidemiologic studies using standardized methods to measure BP and assessing a wider range of variables related to BP. Also,
BP trends should be analyzed by using a life-course perspective\(^8\) (eg, by also considering preconceptional parental characteristics, pregnancy course, early life characteristics [eg, birth weight], nutrition, and physical activity during childhood and adolescence, socio-economic position at different life stages, etc). The lack of firm evidence linking the epidemic of obesity to the prevalence of elevated BP in children, however, is no reason to weaken efforts to curb the current pediatric epidemic of obesity in view of the numerous detrimental consequences of overweight on health in children.\(^{13,62–64}\)

**ACKNOWLEDGMENT**

This work was funded partially by Swiss National Science Foundation grant 3200B0-109999/1.

**REFERENCES**


STUDIES COUNT COSTS OF BIRTH DEFECTS

“Birth defects lead to more than $2.5 billion a year in hospital costs, according to the first national studies to estimate their financial burden on American families. One study, by researchers at the University of Arkansas and the Centers for Disease Control and Prevention, looked at what hospitals charge—not the actual cost of care. Some conditions were particularly deadly. For example, about 85 percent of babies born with anencephaly, that is, without all or most of their brain and skull, had limited treatment options and died within two days. The average bill for one of these cases was $3800. The most expensive condition was hypoplastic left heart, in which most or all of the two left chambers is missing. Treatment cost about $200 000, on average. A second study, by the Agency for Healthcare Research and Quality, estimated what it cost hospitals to care for patients of all ages with birth defects, which it reasoned was about 40 percent of what the hospitals charged. It found that the average age of patients was 17.5 years, the average hospital stay six days, and the average per-stay cost was $18 600. The aggregated cost for all these visits was more than $2.5 billion.”


*Noted by JFL, MD*
The Late Effects of Childhood Cancer Therapy

Joseph D. Dickerman, MD

Department of Pediatrics, University of Vermont College of Medicine, Burlington, Vermont

The author has indicated he has no financial relationships relevant to this article to disclose.

ABSTRACT

In this article the difficulties that face survivors of childhood cancer therapy are presented, and the late effects of such therapy, separated into nonmalignant and malignant late effects, are discussed according to organ system. Recommendations for monitoring the late effects are set forth. A table listing radiation-therapy site and chemotherapeutic agents and selected late effects that result from their use is provided. Finally, a brief recommendation regarding the establishment of a late-effects clinic is also presented.
Primary care physicians, a category that includes pediatricians, internists, family practitioners, and obstetricians/gynecologists, are and will be the health care providers for childhood cancer survivors. Therefore, it is imperative that they familiarize themselves with the late effects of cancer therapy to prevent or diminish them. Chemotherapy used to treat adult cancers but not those that occur in childhood will not be addressed, and the late effects of surgical therapy will not be discussed. In the last 5 years there has been only 1 comprehensive review in the primary care literature that dealt with the late effects of childhood cancer therapy.1 Since the publication of that article in 2002, a prodigious amount of information on this topic has become available.

The Problem

Ten million individuals in the United States are living with a cancer diagnosis today, 3 times the number of survivors in 1971.2 The 5-year survival rate for adults with cancer is 60% to 65%, and for children it is 80% to 85%.3 In the near future 1 of every 450 individuals in the population will be a long-term survivor of childhood cancer; presently 1 in 640 individuals between 20 and 39 years of age is a childhood cancer survivor.4

The long-term morbidity of childhood cancer survivors, now numbering ~270 000 in the United States, was determined in a landmark Childhood Cancer Survivor Study (CCSS).5 Patients (10 397) treated from 1970 to 1986 with a mean age of 26.6 years (range: 18–48 years) were interviewed and compared with their siblings. At least 1 chronic condition was present in 62.3%, and 27.5% had a severe life-threatening condition. The relative risk (RR) of a chronic condition in a survivor compared with a sibling was 3.3, and for a severe or life-threatening condition the RR was 8.2 compared with a sibling. The cumulative incidence of a chronic health condition 30 years after diagnosis was 73.4%, with a cumulative incidence of 42.4% for a severe, disabling, or life-threatening condition or death resulting from a chronic condition. Two or more chronic health conditions were seen in 37.6% of patients compared with 13.1% in siblings and 3 or more chronic health conditions were seen in 23.8% of patients compared with 5.4% in siblings. The RR of selected severe or life-threatening or disabling health conditions were as follows: major joint replacement if not part of therapy (54), congestive heart failure (15.1), second cancers excluding basal cell and squamous cell carcinomas (14.8), severe cognitive dysfunction (10.5), coronary artery disease (10.4), cerebrovascular accident (9.3), renal failure or dialysis (8.9), hearing loss uncorrected by hearing aid (6.3), legally blind loss of an eye (5.8), and ovarian failure (3.5). Survivors of bone or central nervous system (CNS) tumors and patients with Hodgkin’s lymphoma were at highest risk for these severe or life-threatening conditions and were also more likely to have multiple conditions.

Bone cancer survivors more frequently had severe musculoskeletal problems, congestive heart failure, and loss of hearing, whereas CNS cancer survivors were more likely to have cognitive dysfunction, seizures, and endocrinopathies. Hodgkin’s lymphoma survivors were more likely to have coronary artery disease, cerebrovascular accidents, valvular heart disease, cardiomyopathy, second cancers (breast cancer [BC] in women), and lung and thyroid disease. Exposure to 1 of 5 treatment combinations was associated with a risk of having a severe or life-threatening or disabling condition that was 10 times the expected risk: (1) chest radiation and bleomycin, (2) chest radiation and anthracycline, (3) chest radiation and abdominal or pelvic irradiation, (4) anthracycline and an alkylating agent, and (5) abdominal or pelvic irradiation and an alkylating agent. Adverse psychosocial outcomes such as depression were not included in this analysis.

Approximately 95% of children aged 0 to 14 years are treated in Children’s Oncology Group (COG) centers, and 65% are entered onto clinical trials.7 In another CCSS report to determine the type of medical care received over a 2-year period, 9434 patients were surveyed at a mean age of 26.8 years (range: 18–48 years). Eighty-seven percent reported general medical contact, 71.4% a general physical examination, 41.9% a cancer-related visit, and only 19.2% a visit to a cancer center.8

The advantages and problems of long-term follow-up clinics have been enumerated recently. In a survey sent to directors of 24 comprehensive long-term follow-up programs for pediatric cancer survivors in the United States and Canada, the following were listed as primary benefits: health care provided by clinicians familiar with long-term risks, provision of risk-based screening and surveillance, and targeted education for risk reduction and healthy lifestyles. Barriers to clinic functioning were inadequate resources and finances, low institutional commitment, lack of capacity to care for a growing population, difficulty with ongoing communication with community physicians, and lack of interest and awareness among survivors.9

In a study from Sweden, 335 childhood survivors of acute leukemia, lymphoma, or Wilms’ tumor who were over 18 years of age were sent questionnaires 5 years or more after completion of therapy (response rate: 73%).10 Sixty percent had no regular follow-up visits, and 42% of these patients reported that they missed not having one. One third of the respondents were dissatisfied with the follow-up program, but only 3% with regular follow-up visits found them unnecessary. Complaints subjectively related to their diseases or treatment were reported by 47%. Thirty-four percent of all respondents did not miss having a regular follow-up visit, and neither perceived disease-related complaints nor radiation ther-
apy was a predictor for having a scheduled follow-up visit.

In a report on 650 survivors of childhood cancer, it was determined that ~40% had endocrine problems followed, in order, by hearing or vision difficulties, neurocognitive impairment, cardiopulmonary dysfunction, gastrointestinal disorders, second malignancies, and miscellaneous complications, whereas ~30% had no problems. In another study of 290 survivors of childhood cancer, 41% had endocrine problems, 26% developed organ toxicity, 17% had mobility problems, 15% demonstrated neuropsychological difficulties, 14% were infertile, 13% had vision or auditory difficulties, and 10% had cosmetic problems. Furthermore, pediatric cancer survivors face insurance-coverage difficulties and are at risk for unemployment, especially if they had CNS or brain tumors.

All monitoring recommendations that follow are my own except as noted. The COG recommendations may be found on their Web site.

NONMALIGNANT LATE EFFECTS OF CANCER THERAPY

Cancer therapy may adversely affect the following organs.

Heart

Cardiomyopathy, subclinical left ventricular dysfunction, valvular disease, pericardial disease, and arrhythmias have been reported. Cardiac damage may be caused by anthracyclines (doxorubicin, daunorubicin, mitoxantrone, epirubicin, and idarubicin), and ~60% of patients with childhood cancer are currently being treated with anthracyclines. There may be an increased risk of cardiac damage after anthracycline therapy for pregnant women (the 40% increase in blood volume that occurs during pregnancy may adversely affect the subclinical cardiomyopathy attributable anthracyclines, if present, as well as the cardiomyopathy of pregnancy, if that occurs), females, patients treated at <4 years of age, patients treated with other drugs that affect the heart, or patients for whom thoracic radiation is used.

After anthracycline therapy, the risk of congestive heart failure is 0% to 16%, and for subclinical cardiomyopathy it is 0% to 57%. The risk of cardiac damage, most importantly, depends on the cumulative dose of anthracycline. After more than 6 years of follow-up, 57% of children with acute lymphocytic leukemia treated with doxorubicin 45 to 500 mg/m² (mean: 360 mg/m²) had abnormalities of left ventricular afterload or contractility. However, 17% of patients who received only 1 dose of doxorubicin (45 mg/m²) developed elevated afterload, and patients who received a cumulative doxorubicin dose of only 228 mg/m² had either increased afterload (59%), decreased contractility (23%), or both. Protection against anthracycline-induced cardiac toxicity resulting from free-radical damage may be afforded by dexrazoxane, a free-oxygen-radical scavenger. Other antineoplastic drugs may also cause cardiac damage, but to a lesser degree. Thoracic radiation may also be responsible for cardiotoxicity. The pathophysiology of both anthracycline- and radiation-associated cardiomyopathy has been described.

Monitor the patient with a history and physical examination (H&P) and echocardiogram every year for 2 years and then every 2 years if normal and associated with a normal H&P. If the patient is pregnant, obtain an echocardiogram every trimester with at least 1 evaluation by a cardiologist if she has an abnormal echocardiogram.

Vasculature

Arterial damage resulting in stroke or myocardial infarction has been reported. Patients who receive cranial or thoracic radiation are at risk for these complications. In 1 study of childhood CNS malignancies, stroke occurred in 1.6% of the patients. A CCSS report identified 37 of 4828 leukemia survivors (mean age at diagnosis: 5.9 years) and 63 of 1871 brain tumor survivors (mean age at diagnosis: 7.7 years) diagnosed between 1970 and 1986 who developed late-occurring stroke, defined as presenting 5 or more years after the primary diagnosis. The mean interval from the first cancer diagnosis to late-occurring stroke was 9.8 years for the former and 13.9 years for the latter. When compared with siblings, the RR for late-occurring stroke was 6.4 for leukemia survivors and 29 for brain tumor survivors. A mean dose of cranial radiation therapy of ≥30 Gy was associated, in a dose-dependent fashion, with an increased risk for both groups of survivors. In a pilot study, anthracyclines were shown to cause impaired endothelial function, which suggests that they may play a role in the progression of coronary disease. Venous damage is rare.

Monitor the patient with a yearly H&P.

Lung

Pulmonary fibrosis, restrictive-obstructive lung disease, and delayed interstitial pneumonia have been reported. The risk of these complications increases the younger the age at diagnosis. Pulmonary damage may be caused by thoracic radiation, which is dose and fractionation dependent, and significant abnormalities in pulmonary function have been observed after lung irradiation. Thoracic radiation when combined with bleomycin, actinomycin D, cyclophosphamide, vincristine, or Adriamycin can produce radiation pneumonitis at much lower radiation doses. This synergistic effect is observed in abdominal organs when these drugs are given with abdominal radiation therapy.

A CCSS report on 12 390 children 5 years postdiagnosis demonstrated a significant association between lung radiation and lung fibrosis (RR: 4.3), supplemental oxygen use (RR: 1.8), emphysema (RR: 2), recurrent
pneumonia (RR: 2.2), chronic cough and shortness of breath for >1 month (RR: 2), exercise-induced shortness of breath (RR: 1.8), and abnormal chest wall development (RR: 5), as well as a significant association of lung fibrosis with the carmustine (RR: 1.4), bleomycin (RR: 1.7), busulfan (RR: 3.2), lomustine (RR: 2.1), and cyclophosphamide (RR: 1.6). Chest radiation was associated with a 3.5% cumulative incidence of lung fibrosis 20 years after diagnosis.

If the dose of bleomycin is <450 mg, 3% to 5% of patients are affected, and with >500 mg, 20% of patients affected; with carmustine, up to 30% of patients develop pulmonary fibrosis with doses between 80 and 240 mg/m² given every 6 to 8 weeks for >2 years (cumulative dose: 700–1800 mg/m²), and there is a marked increase in incidence at >1500 mg/m². Methotrexate may cause pulmonary damage, whereas other chemotherapeutic agents not previously discussed are rarely responsible for pulmonary toxicity.

Monitor the patient with an H&P and a pulmonary-function study every other year, unless the results are abnormal, and then every year. Obtain a chest computed tomography scan if there is a significantly abnormal pulmonary-function study result and/or clinical symptoms are present.

Gastrointestinal Tract
Fibrosis leading to partial or complete obstruction may occur as a result of radiation.22

Monitor the patient with an H&P yearly.

Spleen
Functional (radiation-induced) or anatomic (staging splenectomy, which is rarely, if ever, done today) asplenism may predispose the patient to sepsis, a lifelong problem.33

Prevent asplenia with immunizations against Haemophilus influenzae, Streptococcus pneumoniae, and Neisseria meningitidis if not previously immunized and counsel the patient to immediately report a fever or feeling ill to his or her physician. There is an ongoing debate regarding the use of prophylactic antibiotics.34 The Working Party of the British Committee for Standards in Clinical Haematology Task Force recommends, as do I, lifelong antibiotic prophylaxis, whereas the American Academy of Pediatrics states that such prophylaxis may be discontinued after 5 years of age.35

Liver
Hepatitis and cirrhosis can occur. Liver damage may be caused by methotrexate and 6-mercaptopurine as well as many other agents and contaminated blood products.40–42 Hepatitis B and C infections secondary to transfusion therapy are rarely seen today; however, for patients treated in the past this may still be a significant problem, especially the increased risk of hepatic cancer secondary to hepatitis C infection.

Monitor the patient with an annual H&P, and check aspartate aminotransferase and alanine aminotransferase levels yearly.

Kidney
Nephropathy, both glomerular and/or tubular damage, can occur.43 Kidney damage may be caused by radiation44 and ifosfamide as well as other antineoplastic drugs such as cisplatin.45

Monitor the patient with H&P, urinalysis, microscopic examination of the urine, complete metabolic panel (including magnesium), and blood pressure assessment yearly.

Bladder
Hematuria, cystitis, fibrosis, and dysfunctional voiding can occur. Bladder damage may be caused by radiation, cyclophosphamide, and ifosfamide. The metabolic by-product of these 2 drugs is acrolein, which irritates the bladder, and the incidence of the adverse effects caused by cyclophosphamide and ifosfamide can be decreased with hydration and mesna, which binds acrolein. The incidence of bladder damage is 5% to 10% for cyclophosphamide46 and 20% to 40% for ifosfamide.49

Monitor the patient with H&P, urinalysis, and microscopic examination of the urine yearly.

Skeletal
Osteopenia, osteoporosis, avascular necrosis, spinal deformities, and other skeletal changes can occur. Skeletal damage may be caused by steroids, methotrexate, cranial radiation (decreased growth hormone with resultant abnormal bone metabolism), direct radiation to the bone, and cyclophosphamide/ifosfamide (gonadal damage leading to ovarian and/or Leydig cell dysfunction with resultant loss of bone mass).51,52

In patients with childhood acute lymphocytic leukemia, there may be a decrease in bone mineral density, the severity of which decreases with time after treatment.53 The significance of this finding in these patients is presently unclear.54 Bone mineral density has also been reported to be reduced in up to one third of survivors of childhood brain tumors, and the reasons are multifactorial, with craniospinal irradiation probably being the most important factor.55 The spinal deformities and other skeletal changes that may result after radiation therapy are seen less frequently now because of lower doses and newer radiation techniques.

Monitor the patient with a yearly H&P including scoliosis screening, and perform a bone density study once 2 or 3 years after therapy; if results are normal, it does not need to be repeated unless there are clinical symptoms and/or signs that suggest a problem.
Muscle
Atrophy may occur after direct radiation to the muscle.66 Monitor the patient with an H&P yearly.

Thyroid
Hypothyroidism or hyperthyroidism may occur after thoracic, cranial, or neck radiation. A CCSS study of 13,674 patients with Hodgkin’s lymphoma57 found an increased risk of hypothyroidism with an increased dose of radiation (>4.5 Gy), with older age (>15 years), with female gender, and <5 years after diagnosis. The RR for hypothyroidism in this study was 17.1, and it occurred in 25% of patients (50% at 20 years if >4.5 Gy was given). It developed at a mean of 7 years after diagnosis, and the median age at treatment was 14 years. The RR for hyperthyroidism was 8, and it occurred in 5% of the patients, with a mean age to development of 8 years.

In a study of 461 children treated for Hodgkin’s lymphoma,58 43% developed hypothyroidism (47% of white patients and 21% of black patients) at a median of 2.9 years after therapy, and the risk was greater for female patients. In an older study of 1,787 patients (~35% younger than 22 years) with Hodgkin’s lymphoma,59 overt and subclinical hypothyroidism was seen in 44% of the patients 20 years after therapy with >30 Gy and in 27% of the patients 20 years after therapy with 7.5 to 30 Gy. Most cases were identified during the second or third year after therapy. The majority of cases occurred in patients treated at 15 to 25 years of age, and the risk was increased in female patients. Hyperthyroidism was seen in 1.7% of the patients, and the risk was 7.2 to 20.4 times that subjects without Hodgkin’s lymphoma.

Monitor the patient with an annual H&P, and check thyrotropin and free-thyroxine levels yearly.

Growth and Development
Obesity
The prevalence of obesity after therapy for childhood acute lymphocytic leukemia is 16% to 56% and is caused by cranial radiation (growth hormone deficiency), steroid therapy, physical inactivity, and increased dietary intake.60 In adult survivors of childhood acute lymphocytic leukemia, cranial radiation at >20 Gy has been associated with obesity, particularly in girls who were treated at 0 to 4 years of age.61 Chemotherapy without cranial radiation may also lead to obesity in survivors of childhood acute lymphocytic leukemia.62 Monitor the patient with an H&P (weight determination) and dietary counseling yearly.

Short Stature
In patients with childhood acute lymphocytic leukemia, short stature may be caused by cranial radiation because of growth hormone deficiency,63 and there is an increased risk if the patient is <4 years of age. Early growth deceleration with bone age retardation is seen, and at the end of therapy ~70% of patients will show a variable degree of catch-up growth; this catch-up growth can be complete 2 to 3 years after treatment in patients who did not receive cranial radiation, but it is usually incomplete in patients who did receive cranial radiation. Because growth is impaired in patients with acute lymphocytic leukemia who did not receive cranial radiation, chemotherapy and/or other factors may also be responsible for this problem.64 Maximizing final height with growth hormone treatment may be achieved if therapy is initiated at the earliest bone age that is clinically feasible.65 Monitor the patient with an H&P (height measurement) yearly and bone age films when clinically indicated.

Gonads
Gonadal failure can occur. Testicular or ovarian damage may be caused by radiation therapy (directly to the gonads or to the brain [hypothalamic-pituitary axis damage]) or alkylating agents (cyclophosphamide, nitrosourea, chlorambucil, ifosfamide, dacarbazine, thiopeta, melphalan, busulphan, carmustine, lomustine, cytarine, or procarbazine). Sertoli cell function (sperm production) is impaired at lower drug doses when compared with impairment of Leydig cell function (testosterone production). The degree of gonadal impairment is related to the age at, dose of, and fractionation schedule for radiation therapy and the age at and dose of chemotherapy at the time of treatment.30,66,67 Prepubertal and adolescent girls are more resistant to alkylator-induced and radiation-induced failure because of increased numbers of follicles. Most young female patients treated with standard chemotherapy will retain ovarian function. The ovaries of younger female patients are more resistant to radiation injury than older ones, but >20 Gy produces failure in most female children. In prepubertal girls, >20 to 30 Gy may lead to failure or incomplete pubertal development. In boys, alkylator-induced Leydig cell failure requiring testosterone replacement is uncommon; however, germ cell damage and infertility are common. Gonadal failure is dose and fractionation dependent with radiation therapy; >3 Gy usually produces irreversible azoospermia. With <12 Gy, Leydig function is usually spared in prepubertal boys.66,67 Fertility may be preserved with sperm banking,68 ovarian transposition, or egg or ovarian tissue banking.69,70 Monitor the patient with an H&P (height measurement, weight determination, and Tanner assessment) and check follicle-stimulating hormone, luteinizing hormone, testosterone, and estradiol levels when clinically indicated. Follicle-stimulating hormone and luteinizing hormone levels are high when direct gonadal radiation
and/or alkylator therapy is responsible for the gonadal failure and low when cranial radiation is responsible.

Central Nervous System

Intellectual or cognitive impairment involving mental processing, attentional or memory deficits, visual-spatial abilities, attention concentration, nonverbal memory, and somatosensory function may be caused by cranial radiation. There is an increase in impairment associated with female gender, decreasing age at therapy (<4 years), and increase in radiation dose. A recent study demonstrated that intrathecal methotrexate combined with either high-dose or very high-dose intravenous methotrexate did not result in poorer neurocognitive, cognitive, or academic outcomes in children treated for acute lymphocytic leukemia when compared with population norms.

Monitor the patient with an H&P yearly and neuro-psychological testing when clinically indicated.

Psychological maladjustment, mood disturbances, behavioral problems, somatic distress, academic under-achievement, unemployment, and/or posttraumatic stress disorder may occur in 10% to 20% of long-term survivors of childhood cancer. In a study of 9535 patients from the CCSS group, 44% reported at least 1 recommendation. Exercise, and/or failure to adhere to sun-protection recommendations may lead to multiple health-compromising behaviors such as smoking, lack of exercise, and/or failure to adhere to sun-protection recommendations.

Monitor the patient with an H&P yearly and neuro-psychological testing when clinically indicated.

Monitor the patient with an H&P yearly and neuro-psychological testing when clinically indicated.

Neuropathy may be caused by vincristine and can lead to impaired deep-tendon reflexes (most common), paresthesias, sensory symptoms, motor weakness, paralytic ileus, and cranial neuropathies. This neuropathy may be enhanced by the use of concurrent itraconazole. Thalidomide may cause a sensory neuropathy, and up to 50% of patients do not recover, whereas the sensory neuropathy caused by cisplatin may persist and can affect up to 20% to 60% of patients.

Monitor the patient with an H&P (with a neurologic examination) yearly.

Leukoencephalopathy, defined as demyelination, white matter necrosis, calcification, and glial damage, may develop after CNS radiation and/or intrathecal methotrexate. It presents with seizures, focal motor signs, dementia, ataxia, and cognitive abnormalities. The risk correlates with methotrexate dose and route, age at treatment, and amount of radiation. A computed tomography scan of the brain may be used to diagnose this condition.

Eye

Cataracts may be observed and can be caused by radiation, steroids, and busulphan. Cranial radiation may result in keratoconjunctivitis.

Monitor the patient with an H&P (with a lens examination) yearly.

Ear

Sensorineural, high-frequency hearing loss may develop after cisplatin therapy, most often observed at a cumulative dose approaching 400 mg/m². Ifosfamide and cranial radiation may exacerbate cisplatin-related hearing loss. New therapies for preventing cisplatin-induced ototoxicity are in the early stages of development. It is thought that cisplatin suppresses the formation of endogenous antioxidants that protect the inner ear against reactive oxygen species, and these new therapies are free-radical scavengers.

Monitor the patient with a yearly H&P and hearing testing. Patients should be carefully evaluated for speech delay or abnormality as well as school performance, because either may be the first manifestation of hearing loss.
Teeth and Gums
Defective dentition, increased caries, root abnormalities, and periodontal disease can occur. These oral abnormalities may be caused by radiation to the area, and there is an increased risk if >2.5 Gy is used or the patient is <5 years of age. Even 0.4 Gy may cause some problems. Chemotherapy may also cause dental defects.85,86
Monitor the patient with a yearly H&P and dental examination by a dentist.

SECOND MALIGNANT NEOPLASMS RESULTING FROM CANCER THERAPY
Childhood cancer survivors are at >19-fold increased risk for developing another malignancy.87 The CCSS reported 314 second malignant neoplasms (SMNs) in 13 518 patients, which yields a standardized incidence ratio (SIR) of 6.38, with the largest excess observed for bone cancer (SIR: 19.14) and BC (SIR: 16.18). An increased risk for SMNs was seen in females and those who were at a younger age when diagnosed. Twenty years after cancer diagnosis, the cumulative estimated SMN incidence was 3.2%.88

Skin Cancer
Skin cancer is probably the result of decreased immunity coupled with sun exposure. Malignant melanoma and nonmelanotic skin cancers represent 10% to 20% of second malignancies after cancer therapy, and there is an increased risk of developing skin cancer over time. There is also an increased incidence of melanocytic nevi, a risk factor for malignant melanoma, in uncommon sites such as the palms and soles of children receiving chemotherapy.89

A CCSS study of 13 132 patients reported 213 patients (1.6%) with nonmelanotic skin cancers, which accounted for the most commonly observed (41%) subsequent cancer.90 Forty-six percent of the patients had multiple occurrences, 90% had received radiation, and 90% of the cancers were in the radiation field. The median age of occurrence was 31 years (range: 7–46 years). Radiation therapy was associated with a 6.3-fold increased risk of skin cancer.
Monitor the patient with an H&P and prevent skin cancer by advising the patient to avoid the sun, wear a hat, cover the skin, and have a yearly skin examination by a dermatologist.

Breast Cancer
BC may develop after thoracic radiation, usually after treatment for Hodgkin’s lymphoma. There are ~120 000 survivors of Hodgkin’s lymphoma in the United States, and SMNs are the leading cause of death for long-term survivors of Hodgkin’s lymphoma, with BC the most frequent solid tumor in women who were treated for Hodgkin’s lymphoma.91 The annual excess incidence of BC increases as the age of patients who are post-Hodgkin’s lymphoma therapy increases. The risk of BC after thoracic radiation for Hodgkin’s lymphoma in women <30 years of age is elevated 4- to 56-fold92,93 depending on the dose of radiation and the age at treatment, with the highest risk in women who were treated at 10 to 20 years of age.92–98 The risk of BC increases with increasing dose of radiation, with each gray unit received by any breast increasing the excess RR of BC by 0.13.96

In a study of adult survivors of pediatric Hodgkin’s lymphoma from the Late Effects Study Group, no increased risk of BC was found if <26 Gy was given,93 and current protocols do not usually use doses greater than this. The risk of BC in a patient >30 years of age at diagnosis of Hodgkin’s lymphoma declined significantly, and there was almost no increased risk of BC in women who were treated after they were 30 to 40 years of age.98,99 The median time to development of BC is 15 years after therapy,94 and hormonal stimulation may be necessary for the development of radiation-induced BC, because 1 study showed a decreased risk of BC associated with ovarian damage that resulted from alkylating agents or radiation.91

Most secondary BCs occur in the field of radiation and are invasive ductal carcinomas, although ~13% are ductal carcinoma in situ. There is an increased incidence of BC in the contralateral breast.100 Whether there is an increased incidence of BC in males who receive thoracic radiation has not been determined yet.

Monitor the patient by having her perform monthly breast self-examination, have an H&P yearly, and, in addition, obtain a yearly mammogram beginning 5 years after diagnosis; this recommendation is endorsed by others,100 because a number of BCs after therapy for Hodgkin’s lymphoma have occurred as early as 5 years after therapy. The COG recommends a mammogram 8 years after therapy or at age 25, whichever is later.15 Breast MRI is more sensitive for detection of BC than ultrasound or mammography in women who are at increased risks of BC secondary to BRCA gene mutations.101 Although it is not clear whether this particular surveillance method will result in reduced mortality, I believe that breast MRI rather than mammography should be used in patients who received thoracic radiation, especially those who were treated when radiation doses were larger and delivery was not as refined.

Thyroid Cancer
Thyroid cancer (TC) may develop after head, neck, or thoracic radiation. In a CCSS study of 14 054 patients, of whom 69 developed TC,102 the incidence of TC in patients with any first malignancy who had radiation therapy to the head, neck, or thorax increased with an increasing dose of radiation. There was an increased...
incidence at 20 to 29 Gy but a decrease in incidence if >30 Gy (ie, killing effect). The estimated RR of TC was 1.32/Gy for a dose <15 Gy. The increased or decreased risk of TC was more pronounced in patients who were diagnosed with a first primary malignant disease before the age of 10 years as opposed to those patients who were older than 10 years. Chemotherapy had no effect on the development of secondary TC. In this study, patients younger than 10 years of age had a substantially higher TC risk over the entire radiation-dose range then did those patients who were older than 10 years, which demonstrated the increased susceptibility of the thyroid gland in young patients. Of the 69 patients with TC, 42% had a first diagnosis of Hodgkin’s lymphoma and 20% had a first diagnosis of leukemia. Other studies have demonstrated a similar increased risk for TC after radiotherapy (RR: 18.3%; RR: 15.6%; absolute excess risk [AER]: 1.4%; SIR: 36.4%; SIR: 35 at 0.5 Gy and 73 at 3.6 Gy).

A thyroid nodule can be palpated in 4% to 7% of adults, and the incidence of nodules is increased if ultrasound is used for detection or the presence of nodules is determined at autopsy. At a mean of 11 years after therapy for Hodgkin’s lymphoma, up to 44% of childhood cancer survivors who received head or neck radiation had detection of nodules when screened by ultrasound. In a report of 647 pediatric patients treated for Hodgkin’s lymphoma, it was determined that 67 (10.4%) of these patients developed 1 or more nodules. Four patients had thyroid disease before diagnosis, and 19 (28%) had received thyroid hormone replacement therapy before the diagnosis of a nodule. Seven patients (10%) had TC diagnosed at a median of 16.2 years after therapy for Hodgkin’s lymphoma (range: 8.4–23.7 years). Only 1 TC was found by ultrasound, and the rest were found by palpation or clinical symptoms (53 of 67 had ultrasound). Forty-one of the 67 patients had asymptomatic nodules detected only by ultrasound. Thirty-four of the patients had a clinical course and imaging characteristics that were consistent with benign nodules and had no biopsy. The median size of the nodules was 0.7 cm (range: 0.2–2.0 cm), and 30% resolved. The majority of TCs resulting from radiotherapy are papillary, and they have a low mortality rate even after metastases.

Monitor the patient with an H&P and testing of thyroid function annually. Obtain a fine-needle aspiration if the nodule is enlarging or is >10 mm or if hypoechoic areas are seen on ultrasound.

Leukemia

Treatment-related acute myelogenous leukemia (AML)/myelodysplastic syndrome (MDS), known as t-AML/MDS, may be caused by alkylating agents (which are also linked to bone and bladder cancers) and topoisomerase 2 inhibitors (etoposide [VP-16] and teniposide [VM-26]). The long-term survival rate of pediatric patients with t-AML/MDS treated with allogeneic stem cell transplantation is 15% to 24%. t-AML/MDS is one of the few late effects for which the incidence plateaus or reverts to normal after a period of time.

With alkylating agents, the risk of t-AML/MDS increases with increasing dose of and age at treatment. The incidence of t-AML/MDS is 0.8% to 2.8%, and the median latency period is 4 to 6 years (range: 1–20 years). There is usually a loss or deletion of chromosome 5 or 7. With topoisomerase 2 inhibitors, there is an increased risk of t-AML (M4 or M5 morphology)/MDS with increasing dose, and it usually occurs in younger patients. The cumulative risk is 0.5% to 18.4%, depending on the dose intensity, and the median latency period is 1 to 3 years (range: 0.5–4.5 years). There is usually rearrangement involving the MLL gene on chromosome band 11q23.

The short-term use of granulocyte colony-stimulating factor after etoposide therapy may increase the risk of developing t-AML/MDS. Although the risk of t-AML/MDS after the use of anthracyclines is thought to be relatively low, 1 study suggests otherwise. Survivors of pediatric Hodgkin’s lymphoma have a 4 to 175 times increased risk for developing t-AML/MDS. The AER of t-AML/MDS was 6.3 in a study of 32 591 adult patients, whereas the SIR was 174.8 in a Late Effects Study Group report of 1380 pediatric patients with Hodgkin’s lymphoma.

Monitor the patient with an H&P and complete blood count yearly. Mildly abnormal complete blood count values (low white blood count, platelets, and hemoglobin and an elevated mean corpuscular volume) are common in survivors of childhood cancer; they tend to persist, are of questionable significance, and do not seem to reflect preleukemic changes. Telomere shortening resulting from chemotherapy, which leads to repeated cycles of hematopoietic regeneration, is thought to be responsible for the hematopoietic stem cell injury.

Sarcoma of Bone and Connective Tissue

Bone and soft tissue sarcomas may occur after radiation therapy, and the risk is proportional to the dose and the concurrent use of alkylating agents. In a report from the British Childhood Survivors of Cancer Study of 13 175 patients diagnosed between 1940 and 1983, the cumulative probability of developing bone cancer after radiation therapy for the entire cohort was 0.9% within 20 years after treatment. The risk was increased to 7.2%
after hereditary retinoblastoma (a malignancy that predisposes a patient to the development of osteosarcoma), 5.4% after Ewing sarcoma, and 2.4% after other malignant bone cancers. The RR of sarcomas, with bone and connective tissue being the most common sarcomas, after pediatric Hodgkin’s lymphoma therapy was reported to be 1.3 to 37.1.97 In another study, the AER for bone/connective tissue cancers after therapy for Hodgkin’s lymphoma was 2.3 in patients <21 years of age, and the AER declined in patients who were older when treated.12 In a report of 11 183 patients <21 years of age,120 Ewing sarcoma–family tumors rarely occurred after treatment of primary childhood cancer (1.3% of 479 second cancers). Most of these tumors did not seem to be related to radiation therapy, and long-term survival was possible.

Monitor the patient with an H&P yearly.

Second Carcinomas Other Than Breast, Thyroid, and Skin

The CCSS group, in the largest study to date of secondary malignancies and all SMNs resulting from treatment of brain tumors, identified 71 carcinomas (excluding breast, thyroid, and skin) in 13 136 patients diagnosed from 1970 to 1986 at <21 years of age.121 The overall SIR for a subsequent adult-type carcinoma was 4 (a fourfold increased risk) and was significantly elevated for all primary cancer diagnoses except CNS neoplasms. Patients with SMNs were more likely to be older and to have a primary diagnosis of Hodgkin’s lymphoma, soft tissue sarcoma, or neuroblastoma, a history of a first-degree relative with cancer, and a history of alcohol use. The overall cumulative incidence of developing a second carcinoma was 0.45% at 20 years of follow-up. Table 1 lists the type of second carcinomas with the unadjusted SIR for both the risk of the carcinoma and the male and female distribution.

Survivors of neuroblastoma, soft tissue sarcoma, and Wilms’ tumor had the greatest risk of developing a subsequent carcinoma, with SIRs of 24, 6.2, and 4.8, respectively. The second carcinoma occurred at a median age of 27 years (range: 10–44 years), and the median elapsed time between the development of the second carcinoma and the primary therapy was 15 years (range: 6–28 years). In contrast, the incidence of the majority of carcinomas in the general population rises from 41 to 50 years of age and reaches it peak at 50 to 70 years.

Survivors of Wilms’ tumor had an increased risk of developing colorectal (SIR: 25.4) and other gastrointestinal carcinomas (SIR: 18), as did the survivors of Hodgkin’s lymphoma (SIR: 2.5 and 7.4, respectively). For most sites, patients <10 years of age at diagnosis had a greater risk of developing secondary carcinomas. Patients who were treated with topoisomerase 2 inhibitors (etoposide, teniposide) had an increased risk of developing lung cancer (SIR: 73.4 vs 1.8 in nonexposed patients), as did patients treated with alkylating agents (SIR: 7 vs 0 in nonexposed patients). Platinum therapy was associated with an elevated SIR for colorectal (SIR: 14.7) and kidney (SIR: 48.7) carcinomas. Radiation therapy was associated with an increased risk of all secondary carcinomas except for the reproductive organs, and the SIR was most marked for head and neck carcinomas (SIR: 18.5 vs 2.3 in nonirradiated patients). The site of the second carcinoma occurred in the previously irradiated field for all (4 of 4) lung carcinomas, in 85% (17 of 20) for head and neck carcinomas, and in 71% (10 of 14) of gastrointestinal carcinomas.

Of 71 patients, 22 (33%) had second carcinomas in an area that was not exposed to radiotherapy, and 11 (50%) of the 22 had no previous radiation. Sixteen of these patients (73%) received alkylating agents. Four (5.6%) of the 71 patients had received neither chemotherapy nor radiation therapy. The authors of this study cautioned that the sample size was too small to perform adjusted analysis to determine the independent contribution of treatment and patient factors.

In addition to a yearly H&P, monitor for colon cancer in those patients who received abdominal radiation with colonoscopy every 10 years beginning 15 years after therapy or at age 35, whichever is later; this is a COG recommendation.15

<table>
<thead>
<tr>
<th>Type of Second Carcinoma</th>
<th>Unadjusted SIR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
</tr>
<tr>
<td>Head and neck (highest in patients with primary soft tissue sarcoma [SIR: 21.6], leukemia [SIR: 20.9], and neuroblastoma [SIR: 20.9])</td>
<td>11.7</td>
</tr>
<tr>
<td>Kidney</td>
<td>8.5</td>
</tr>
<tr>
<td>Bladder</td>
<td>4.4</td>
</tr>
<tr>
<td>Other gastrointestinal</td>
<td>1.8</td>
</tr>
<tr>
<td>Male genitourinary</td>
<td>3.5</td>
</tr>
<tr>
<td>Lung</td>
<td>46</td>
</tr>
<tr>
<td>Colon and rectum</td>
<td>2.8</td>
</tr>
<tr>
<td>Female genitourinary</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA indicates not applicable.

Brain Tumors Resulting From Treatment of All Childhood Malignancies and All SMNs Resulting From Treatment of Brain Tumors

Brain tumors resulting from cranial radiation have a latency period of 9 to 10 years, and the younger the age at treatment, the greater the risk.122 The CCSS reported 116 subsequent CNS neoplasms in 14 361 5-year survivors of childhood cancers.123 Gliomas (n = 30) occurred at a median of 9 years from the original diagnosis, and the SIR was 8.7. Meningiomas (n = 66) occurred at a median of 17 years from the original diagnosis. Other subsequent CNS tumors were primitive neuroectodermal tumor (n = 6) and CNS lymphoma (n = 1). After adjustment for radiation dose, neither original cancer...
diagnosis nor chemotherapy was associated with an increased risk.

A National Cancer Institutes Surveillance and End Results (SEER) analysis of 2056 survivors of childhood brain tumors revealed that there were 39 SMNs. There was a 4.7-fold increase in risk in patients treated from 1979 to 1984 and a 6.7-fold increase in risk in those treated after 1985; the difference was possibly related to treatment intensity with the increasing use of chemotherapy in addition to the standard radiation therapy. For astrocytomas, the 3 most common SMNs were “other tumors” (5), followed by fibrosarcomas (3) and melanomas (3). For primitive neuroectodermal tumors, the 2 most common SMNs were “other tumors” (3) and acute lymphocytic leukemia (2). For other gliomas, astrocytomas (3) were the most common SMN. The mean age at primary diagnosis was 9.5 years, and there was a median time to SMN diagnosis of 14.1 years.

Monitor the patient with an H&P (with neurologic examination) yearly.

Bone Marrow and Stem Cell Transplantation
In addition to the usual late effects of chemotherapy and radiation therapy, patients posttransplant are also at risk of developing chronic graft-versus-host disease and all its attendant complications. They are also at risk for developing secondary malignancies. In 1 study, the SIR for second malignancies was 5.4 (excluding post-transplant lymphoproliferative disorder, which usually develops within the first year after transplant), and the SIR for t-MDS/AML was 300.

To facilitate the use of the information provided above, I have formulated a table (Table 2) that lists radiation-therapy site and chemotherapeutic agents and selected late effects resulting from their use. The reader may refer back to the section on each individual organ, both in the nonmalignant and malignant late-effects section, for a more detailed discussion of the late effects of cancer therapy on each organ system as well as recommendations for patient monitoring.

**SUMMARY AND CONCLUSIONS**
The number of long-term, adult survivors of childhood cancer therapy will continue to increase, and almost 75% will have a chronic health problem resulting from their cancer therapy, whereas >40% will have a severe, disabling, or life-threatening condition or death caused by a chronic condition that resulted from their therapy.

I believe that in addition to being followed by their primary care physician, all long-term survivors of childhood cancer therapy should attend a specialized late-effects clinic yearly and be evaluated by a member of the oncology team, either the physician or the pediatric oncology nurse practitioner, who initially treated them. A psychologist and social worker should always be present, and subspecialists should be available on or near
the site as needed. Ideally, such clinics should be located in the same center in which the patient was initially treated but scheduled at a different location or time from the regular hematology/oncology clinic. Four possible models of late-effects clinics with their advantages and disadvantages have been described, and a clinic-based comprehensive care model for survivors of childhood cancer has been proposed. Recently, 11 review articles that dealt with a number of issues concerning cancer survivorship (including follow-up, late effects, models for delivering care, patient education, patient advocacy, employment and psychosocial concerns, health promotion, and research directions in survivors of pediatric and adult cancers) have been published.

With no or less radiotherapy being delivered with newer equipment in better fractionation schedules, and the replacement of or the use of reduced doses of second-cancer–inducing chemotherapy, the late effects resulting from such current treatment will decrease when compared with those reported here. In fact, for the most recent protocols for some patients with Hodgkin’s lymphoma, radiation therapy has been omitted or delivered at lower doses, as it has in the treatment of some abdominal neoplasms, whereas in childhood acute lymphocytic leukemia it is used only in high-risk patients. However, there are still many long-term survivors of childhood cancer therapy who were treated 10 to 50 years ago, and these patients need particularly close monitoring.

The risk estimates reported here, which were determined from patients treated a number of years ago, are probably too high for current use, because cancer therapy has changed dramatically in recent years; however, new cancer therapies used now or in the future will, in all likelihood, be associated with their own late effects. Therefore, patients who are treated with these new therapies must be monitored closely to assess the magnitude of any late effects. On the COG Web site there are patient-education handouts that described the late effects of cancer therapy, which should be given to patients during their first visit to the late-effects clinic. For readers interested in the embryonic field dealing with the late effects of cancer therapy in adult survivors of adult cancer, the Institute of Medicine report on this subject can also be found on the Web.

ACKNOWLEDGMENT
I thank Nancy Moreland for invaluable help with the preparation of this article.

REFERENCES
3. Hampton T. Cancer treatment’s trade-off: years of added life can have long-term costs. JAMA. 2005;294:167–168

<table>
<thead>
<tr>
<th>Type of Treatment</th>
<th>Late Effect(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avascular necrosis</td>
<td>Skin</td>
</tr>
<tr>
<td>Spinal deformities</td>
<td>Melanotic nevi</td>
</tr>
<tr>
<td>Cancer (sarcoma)</td>
<td>Nonmelanotic skin cancer</td>
</tr>
<tr>
<td>Any radiation</td>
<td>Skin</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>Bone marrow</td>
</tr>
<tr>
<td>Alkylating agents</td>
<td>Bone</td>
</tr>
<tr>
<td>Nitrosourea</td>
<td>Cancer</td>
</tr>
<tr>
<td>Cytarabine</td>
<td>Bladder</td>
</tr>
<tr>
<td>Cyclophosphamide</td>
<td>Cystitis (cyclophosphamide, ifosfamide)</td>
</tr>
<tr>
<td>Ifosfamide</td>
<td>Fibrosis (cyclophosphamide, ifosfamide)</td>
</tr>
<tr>
<td>Chlorambucil</td>
<td>Dysfunctional voiding (cyclophosphamide, ifosfamide)</td>
</tr>
<tr>
<td>Carmustine</td>
<td>Cancer</td>
</tr>
<tr>
<td>Lomustine</td>
<td>Bladder</td>
</tr>
<tr>
<td>Thiotepa</td>
<td>Cystitis (cyclophosphamide, ifosfamide)</td>
</tr>
<tr>
<td>Procarbazine</td>
<td>Fibrosis (cyclophosphamide, ifosfamide)</td>
</tr>
<tr>
<td>Melphalan</td>
<td>Sensory nerve damage (cyclophosphamide)</td>
</tr>
<tr>
<td>Busulfan</td>
<td>(exacerbated by ifosfamide and cranial radiation)</td>
</tr>
<tr>
<td>Cisplatin</td>
<td>Kidney</td>
</tr>
<tr>
<td>Methotrexate</td>
<td>Nephrotoxicity (ifosfamide)</td>
</tr>
<tr>
<td>6-Mercaptopurine</td>
<td>Eye</td>
</tr>
<tr>
<td>Methotrexate</td>
<td>Cataracts (busulfan)</td>
</tr>
<tr>
<td>Vincristine</td>
<td>Cataracts</td>
</tr>
<tr>
<td>Anthracyclines</td>
<td>Cataracts</td>
</tr>
<tr>
<td>Doxorubicin</td>
<td>Skeletal</td>
</tr>
<tr>
<td>Daunorubicin</td>
<td>Osteoporosis</td>
</tr>
<tr>
<td>Idarubicin</td>
<td>Cardiomyopathy</td>
</tr>
<tr>
<td>Epirubicin</td>
<td>Arrhythmias</td>
</tr>
<tr>
<td>Mitoxantrone</td>
<td>Left ventricular dysfunction</td>
</tr>
<tr>
<td>Bleomycin</td>
<td>Cardiac failure</td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>Interstitial pneumonitis</td>
</tr>
<tr>
<td>Topoisomerase 2 inhibitors</td>
<td>Bone Marrow</td>
</tr>
</tbody>
</table>

43. Smith GR, Thomas PR, Ritchey M, Norkool P. Long-term renal function in patients with irradiated bilateral Wilms tu-


85. Sonis AL. Craniofacial development, teeth and salivary
glands. In: Wallace WH, Green DM, eds. Late Effects of Child-
86. Alberth M, Kovalecz G, Nemes J, et al. Oral health of long-
43:88–90
87. Argani P, Lao M, Ballard ET, et al. Translocation carcinomas of
the kidney after chemotherapy in childhood. J Clin Oncol.
2006;24:1529–1534
neoplasms in five-year survivors of childhood cancer: Child-
hood Cancer Survivor Study. J Natl Cancer Inst. 2001;93:
618–629
WH, Green DM, eds. Late Effects of Childhood Cancer. London,
United Kingdom: Arnold; 2004:321–331
in survivors of childhood and adolescent cancer: a report from
the Childhood Cancer Survivor Study. J Clin Oncol. 2005;23:
3733–3741
radiotherapy and chemotherapy among young women with Hog-
dkin disease [published correction appears in JAMA.
neoplasms among long-term survivors of Hodgkin’s disease: a
population-based evaluation over 25 years. J Clin Oncol. 2002;
20:3484–3494
neoplasms continues with extended follow-up of childhood
Hodgkin’s disease: report from the Late Effects Study Group.
94. Yahalom J. Breast cancer after Hodgkin disease: hope for a
safer cure. JAMA. 2003;290:529–531
95. Wendland MM, Tsodikov A, Glenn MJ, Gaffney DK. Time
interval to the development of breast carcinoma after treat-
96. Hill DA, Gilbert E, Dores GM, et al. Breast cancer risk follow-
ing radiotherapy for Hodgkin lymphoma: modification by other
97. Lin HM, Teitell MA. Second malignancy after treatment of
pediatric Hodgkin disease. J Pediatr Hematol Oncol. 2005;27:
28–36
after radiotherapy for a first cancer during childhood. J Clin
Oncol. 2005;23:197–204
99. Hancock SL, Tucker MA, Hoppe RT. Breast cancer after treat-
100. Wolden SL, Hancock SL, Carlson RW, Goffinet DR, Jeffrey SS,
Hoppe RT. Management of breast cancer after Hodgkin’s dis-
breast ultrasound, and magnetic resonance imaging for sur-
veilance of women at high familial risk for breast cancer.
J Clin Oncol. 2005;23:8469–8476
102. Sigurdson AJ, Ronckers CM, Mertens AC, et al. Primary thy-
orid cancer after a first tumour in childhood (the Childhood
Cancer Cancer Survivor Study): a nested case-control study. Lancet.
103. de Vathaire F, Hardiman C, Shamsaldin A, et al. Thyroid
carcinomas after irradiation for a first cancer during child-
104. Singer PA, Cooper DS, Daniels GH, et al. Treatment guidelines
for patients with thyroid nodules and well-differentiated thy-
1996;156:2163–2172
2004;351:1764–1771
106. Crom DB, Kaste SC, Tubergen DG, Greenwald CA, Sharp GB,
Hudson MM. Ultrasonography for thyroid screening after head
and neck irradiation in childhood cancer survivors. Med
Pediatr Oncol. 1997;28:15–21
107. Metzger ML, Howard SC, Hudson MM, et al. Natural history of
thyroid nodules in survivors of pediatric Hodgkin lymph-
108. Acharya S, Sarafoglou K, LaQuaglia M, et al. Thyroid neo-
plasms after therapeutic radiation for malignancies during
childhood or adolescence. Cancer. 2003;97:2397–2403
carcinoma of childhood and adolescence: a 30-year experi-
ence at the Istituto Nazionale Tumori in Milan. Pediatr Blood
Cancer. 2006;46:300–306
110. Ng A, Taylor GM, Eden OB. Treatment-related leukaemia: a
clinical and scientific challenge. Cancer Treat Rev. 2000;26:
377–391
111. Bhatia S. Epidemiology. In: Wallace WH, Green DM, eds. Late
Effects of Childhood Cancer. London, United Kingdom: Arnold;
2004:57–69
ctic stem cell transplantation for pediatric patients with ther-
apy-related acute myeloid leukemia or myelodysplastic syn-
113. Davies SM. Therapy-related leukemia associated with alky-
114. Libura J, Slater DJ, Felix CA, Richardson C. Therapy-related
acute myeloid leukemia-like MLL rearrangements are induced
by etoposide in primary human CD34+ cells and remain stable
2131
115. Relling MV, Boyett JM, Blanco JG, et al. Granulocyte colony-
stimulating factor and the risk of secondary myeloid malignancy
after etoposide treatment. Blood. 2003;101:
3862–3867
116. Le Deley MC, Leblanc T, Shamsaldin A, et al. Risk of second-
ary leukemia after a solid tumor in childhood according to the
dose of epipodophyllotoxins and anthracyclines: a case-
control study by the Societe Francaise d’Oncologie Pedi-
117. Long ZB, Oeffinger KC, Brooks SL, et al. Incidence and clin-
cal relevance of abnormal complete blood counts in long-
1640
118. Engelhardt M, Ozkaynak MF, Drullinsky P, et al. Telomer-
aise activity and telomere length in pediatric patients with malignancies undergoing chemotherapy. Leukemia. 1998;
12:13–24
alkylation agents, and risk of bone cancer after childhood
120. Spunt SL, Rodriguez-Galindo C, Fuller CE, et al. Ewing sar-
coma-family tumors that arise after treatment of primary
121. Bassal M, Mertens AC, Taylor L, et al. Risk of selected subse-
quent carcinomas in survivors of childhood cancer: a report
from the Childhood Cancer Survivor Study. J Clin Oncol.
2006;24:476–483
122. Bhatia S, Sklar C. Second cancers in survivors of childhood
123. Neglia JP, Robison LL, Stovall M, et al. New primary neo-
plasms of the central nervous system in survivors of child-
hood cancer: a report from the Childhood Cancer Survivor
124. Peterson KM, Shao C, McCarter R, MacDonald TJ, Byrne J.
An analysis of SEER data of increasing risk of secondary

PEDiATRICS Volume 119, Number 3, March 2007 567


---

**CHINA TRIALS**

“In the past few years, foreigners with spinal-cord injuries flocked to China, hoping to find cures unavailable at home. Hundreds of patients have spent millions of dollars to undergo risky surgeries, which usually involve implanting cells culled from aborted fetuses into the spine. Some claim the procedures have helped them regain limited movement. But scientists around the world remain doubtful, given that few, if any, rigorous clinical trials have been conducted. A number of prominent scientists have even challenged the ethics of selling the unproven therapies to patients desperate for a cure to paralysis. Now, Wise Young, a Rutgers University professor and well-known researcher in the field, is trying to change that. He is planning a vast clinical trial—the first and largest of its kind—in hopes of bringing China’s research into spinal-cord injuries from the margins to the medical mainstream. Over the past three years, he has cobbled together a network of about 20 Chinese hospitals to conduct scientific trials that he hopes will pass muster in the West. . . . An estimated 60 000 people suffer spinal-cord injuries each year, compared with about 11 000 new injuries in the US. . . . Dr Young plans to conduct four clinical trials each year over a five-year period at a cost of less than $10 million. His fledgling China Spinal Cord Network has already received approval for its first set of trials, and is awaiting official approval for others.”


Noted by JFL, MD
SPECIAL ARTICLE

Internet-Based Home Monitoring and Education of Children With Asthma Is Comparable to Ideal Office-Based Care: Results of a 1-Year Asthma In-Home Monitoring Trial

Debora S. Chan, PharmD, FASHP, Charles W. Callahan, DO, Virginia B. Hatch-Pigott, MD, Annette Lawless, RN, H. Lorraine Proffitt, RN, Nola E. Manning, RN, Mary Schweikert, RN, Francis J. Malone, MD†

Department of Pediatrics, Tripler Army Medical Center, Honolulu, Hawaii

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

ABSTRACT

OBJECTIVE. The goal was to determine whether home asthma telemonitoring with store-and-forward technology improved outcomes, compared with in-person, office-based visits.

METHODS. A total of 120 patients, 6 to 17 years of age, with persistent asthma were assigned randomly to the office-based or virtual group. The 2 groups followed the same ambulatory clinical pathway for 12 months. Office-based group patients received traditional in-person education and case management. Virtual group patients received computers, Internet connections, and in-home, Internet-based case management and received education through the study Web site. Disease control outcome measures included quality of life, utilization of services, and symptom control.

RESULTS. A total of 120 volunteers (45 female) were enrolled. The groups were clinically comparable (office-based: 22 female/38 male; mean age: 9.0 ± 3.0 years; virtual: 23 female/37 male; mean age: 10.2 ± 3.1 years). Virtual patients had higher metered-dose inhaler with valved holding chamber technique scores than did the office-based group at 52 weeks (94% vs 89%), had greater adherence to daily asthma symptom diary submission (35.4% vs 20.8%), had less participant time (636 vs 713 patient-months), and were older. Caregivers in both groups perceived an increase in quality of life and an increase in asthma knowledge scores from baseline. There were no other differences in therapeutic or disease control outcome measures.

CONCLUSIONS. Virtual group patients achieved excellent asthma therapeutic and disease control outcomes. Compared with those who received standardized office-based care, they were more adherent to diary submission and had better inhaler scores at 52 weeks. Store-and-forward telemedicine technology and case management provide additional tools to assist in the management of children with persistent asthma.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1884
doi:10.1542/peds.2006-1884
† Dr Francis “Buzz” Malone died on January 31, 2005.
The views expressed in this manuscript are those of the authors and do not reflect the official policy or position of the Department of the Army, the Department of Defense, or the US government.

Key Words: telemedicine, home, telehealth, telemonitoring, store-and-forward, clinic, case management, pediatric

Abbreviations: DPI—dry-powder inhaler MDI-VHC—metered-dose inhaler with valved holding chamber

Accepted for publication Oct 30, 2006
Address correspondence to Debora S. Chan, PharmD, FASHP, Department of Pediatrics (MCHK-PE), 1 Jarrett White Rd, Honolulu, HI 96859-5000. E-mail: debora.chan@hawaii.mamc.amedd.army.mil

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275); published in the public domain by the American Academy of Pediatrics
According to the Institute of Medicine, a quality health care system in the United States in the 21st century will be one based on “continuous healing relationships.” In such a system, patient care is customized according to the patients’ needs and values, the patient is the ultimate source of control, and knowledge is shared in a free flow of information. Information technology will be used to change the way care is delivered, from an approach centered on the physician visit to one in which tools such as e-mail and Internet-based health information provide continuous communication and information flow between clinicians and patients. Telemedicine, that is, the exchange of medical information between individuals separated by distance or time, is the term traditionally used for such technology-based medical interventions.

It is not clear what form this continuous flow of information will take or whether outcomes from a telemedicine system based on frequent electronic communication would compare favorably with traditional office-based care of patients with chronic diseases. It is also unclear which patient populations or chronic diseases will be amenable to care with these tools.

Earlier research suggested that children with asthma might be the ideal target population for the use of telemedicine. Asthma has been identified by the Agency for Healthcare Research and Quality as 1 of 15 “priority conditions” for strategic, evidence-based intervention. Traditionally, asthma has been one of the most frequent discharge diagnoses from Department of Defense hospitals, although targeted disease-management programs have made a significant impact on asthma hospitalization rates.

To date, the use of telemedicine in the home monitoring of children with asthma has focused primarily on real-time or “synchronous” connection with medical professionals and exchange of information. Real-time telemedicine systems are inconvenient for patients and providers and depend on temperamental video connections. However, there is increasing experience with Internet-based, “asynchronous,” telemedicine systems for in-home monitoring of patients with asthma, including an initial 6-month pilot study (Telemedicine In-Home Monitoring Evaluation Project) of Internet-based asthma monitoring, which demonstrated that the technology was effective and acceptable to our military patients.

The Asthma In-Home Monitoring trial was a prospective, 1-year, randomized trial of Internet-based, asynchronous, in-home asthma monitoring for children. The trial examined the virtual management of children with asthma with a specially designed Web site that allowed for interaction with trained asthma care managers and asynchronous (“store-and-forward”) collection and review of medical information. The Internet-based care in the Asthma In-Home Monitoring trial was compared with standardized, office-based care following a carefully developed, clinical practice guideline for asthma care.

We hypothesized that asthma in-home monitoring, management, and education with Internet-based store-and-forward technology would lead to improved outcomes, including therapeutic and diagnostic adherence and disease control (quality of life, lung function, utilization of services, rescue therapy, symptom control, patient education, and patient satisfaction), in children with asthma, compared with office-based case management. Furthermore, we hypothesized that asthma in-home monitoring would allow patients to follow the skills, information, and education goals of an ambulatory asthma pathway from their own home, at their own convenience.

**METHODS**

Children with persistent asthma were solicited for enrollment into the study via telephone (with permission from their primary care provider) and on presentation to the pediatric clinic for an asthma visit, from a population of ~40 000 military dependent children on the island of Oahu. The diagnosis of persistent asthma was based on the National Heart, Lung, and Blood Institute Expert Panel Report 2 guidelines published in 1997. Assignment of severity classification was made at the beginning of the study, on the basis of the severity of disease without therapy. In addition to the aforementioned factors, inclusion criteria included the following: dependent of active duty or retired US military personnel, 6 to 17 years of age, not moving from Oahu for 12 months after entry into the study, ability to receive cable modem connections in the home, willing to learn to record and to send inhaler technique and peak flow data 2 times per week, willing to attend asthma education follow-up visits either in person or electronically (virtually) at 2-week, 6-week, 3-month, and 6-month intervals after initiation into the study, willing to complete satisfaction and education surveys at the end of the study period, and willing to sign informed written consent forms. Patients were asked to give written informed consent, and patients >7 years of age were also asked to provide written informed assent. Patients underwent block randomization with a table of random numbers and were enrolled in either the “virtual” group (60 subjects) or the office-based group (60 subjects).

The initial evaluation for all patients included history, physical examination, and pulmonary function tests (spirometry). The research team solicited and enrolled all patients. A single research pediatrician (Dr Hatch-Pigott) assigned the asthma severity classification and determined the appropriate management plan by using National Heart, Lung, and Blood Institute Expert Panel Report 2 criteria, in consultation with a pediatric pulmonologist (Dr Callahan).

Office-based group patients received subsequent
asthma visits in person at the pediatric clinic at Tripler Army Medical Center, with the study pediatrician and 1 of the 4 assigned nurse case managers or the pediatric clinical pharmacist case manager. Patients were treated with an ambulatory asthma clinical pathway, with 6 visits scheduled 0, 2, 6, 12, 26, and 52 weeks after enrollment. At each visit, patients and their parents received in-depth asthma education from the case manager, with specific subjects being determined by an asthma educational pathway. Office-based group patients received all of their information in person at the pediatric clinic. The virtual group received 3 in-person visits, at 0, 26, and 52 weeks, and the rest as virtual visits. Virtual visits included asthma education, a video recording of peak flow meter and inhaler use forwarded to the Web site, daily asthma diaries, and communication with the case manager electronically via the Web site. Both groups had 24-hour/7-day access to their case manager through the Internet (virtual group) and/or telephone (virtual and office-based groups). Data collection for office-based group patients began at the initial visit and ended at the last visit, which was scheduled as close to 12 months as practical for the family.

Patients in the virtual group were provided a home computer system, camera, and Internet access. On-site in-home instruction was provided by technical experts on equipment use and Web site capabilities and use. Each patient received the same models of computer and computer equipment, as well as broadband Internet access. Patients and their parents were taught how to use the equipment and how to record and to submit videos by using a computer-mounted digital video camera, to capture the patient’s peak flow meter and inhaler technique. A customized educational and monitoring Web site was developed, which allowed for secure socket layer interactive asthma education that followed the same curriculum as the office-based asthma education. The site also provided secure e-mail contact between patients and case managers, as well as the capability for digital video uploads. This interaction was deemed compliant with the current Health Insurance Portability and Accountability Act standards. Digital videos of the patients using inhaled medication and the peak flow meter were recorded and loaded to the Web site on a predetermined schedule, according to the protocol.

A detailed asthma symptom diary and quality of life survey were included on the Web site. Patients and families were instructed regarding the submission of daily symptom diary entries. This diary information was entered electronically, directly to the Web site. Videos were recorded and loaded on the site for the case manager, who scored them with standardized checklists and provided instruction through e-mail back to the patient/family. Videos were sent 2 times per week for 6 weeks and then once-weekly thereafter. Data collection for the virtual patients began when they received the computer and initial instruction and ended at the last visit, which was scheduled as close to 12 months as practical for the family.

Patients in both groups were contacted (by telephone for the office-based group and by e-mail for the virtual group) by the case manager 2 times per week for 6 weeks and once per week thereafter, to review the asthma action and home management plans, to assess the symptom diary, to remind the patient to perform and to record peak flow measurements daily in the diary, to remind the patient to complete symptom diary information every day, to answer questions, and to intervene if needed. All patients were able to contact the case manager 24 hours per day, 7 days per week, to obtain an unscheduled “sick” office visit with the pediatrician and the case manager as needed, in addition to their scheduled protocol visits.

A number of outcome parameters were assessed. In broad terms, these included measurements of adherence and disease control. We assessed selected aspects of therapeutic and diagnostic treatment regimen adherence. We defined therapeutic adherence as outcomes that assessed directly adherence to the therapeutic regimens, including controller medication use. All patient prescriptions were filled through military pharmacies by using electronic physician order entry. Computerized records of filled prescriptions were available for comparison and analysis. In addition, the number of completed dry-powder inhaler (DPI) or metered-dose inhaler with valved holding chamber (MDI-VHC) videos was compared with the number expected on the basis of the research protocol. The protocol directed patients to submit videos 2 times per week for 6 weeks and then once per week thereafter.

Diagnostic adherence focused on adherence to instructions designed to assess physiologic disease control. This included daily completion of the asthma symptom diary, electronically by virtual group patients and on paper by office-based group patients. It also included assessment of peak flow meter technique and the use and scoring of peak flow meter data submitted electronically. Virtual group patients were asked to submit digital recordings of their peak flow measurements; their actual number of submissions and their technique were compared with those expected on the basis of the research protocol. The protocol directed patients to submit a video of peak flow measurements 2 times per week for 6 weeks and then once per week thereafter. Peak flow was assessed for office-based group patients during scheduled follow-up visits.

The measures of disease control included lung function tests (spirometry performed at intake and study exit), peak flow (percentage of personal best), patient and caregiver pediatric asthma quality of life questionnaires (analyzed at intake and study exit), utilization of services (emergency department visits, hospitalizations,
and unscheduled acute visits because of asthma, from our centralized medical chart database and case manager records), rescue therapy use (β-receptor agonist use and refills and use of oral prednisone rescue therapy, from computerized pharmacy records), symptom control (diary symptom score), and asthma knowledge retention (preaducation and posteducation testing). The pediatric asthma quality of life questionnaire was used with permission from Dr Elizabeth Juniper. Patient and case manager participation time was also recorded. At the completion of the study, the computers were collected from families in the virtual group.

Special emphasis was placed on the correct use of inhaled medication. MDI-VHC and DPI techniques were standardized and reinforced at every visit. Techniques were assessed at intake and at each visit and were scored with a previously described scale.15 The technique for each group using a MDI-VHC or DPI was compared at intake, at 26 weeks, and at 52 weeks.

The study pediatrician and case manager saw patients in both groups for all scheduled physician visits. The case manager recommended an appointment with the study pediatrician and case manager for patients in either group if one was needed for closer observation or intervention, as determined through telephone or e-mail communication.

Sample size was calculated by using a 2-way analysis of variance with repeated measures over the 3 time points of the pathway visits. Factor A was group, with 2 levels (traditional office-based treatment group and virtual group). Factor B was time period, with 3 levels (baseline, 26 weeks, and 52 weeks). A minimal sample size of 45 patients in each group enabled detection of an effect size of ≥20% at an α level of .05, with 84% statistical power. This sample size of 45 patients in each group would also allow detection of a difference in group means with an effect size of 30%, using a 2-tailed t test, at an α level of .05, with 80% statistical power.

Because of the mobility of our population, which was related to normal military procedures and was accentuated by deployments associated with the war in the Middle East, patient participation was calculated in patient-years. This was derived from the total number of patient days in the study divided by 365 days/year. Computerized data, including history, physical examination, and peak flow monitoring data, were collected and stored on secured government computers. All aspects of the evaluation represented the standard of care for children with asthma.

Comparisons of scores were made by using differences in the individual scores. Statistical significance was calculated by using a paired t test. The statistical significance of differences in continuous outcomes (eg, case manager and participant time) between subjects in the virtual and office-based groups was assessed with unpaired t tests or, if normality was questionable, with the nonparametric Wilcoxon rank test. For comparison of proportions with Fisher’s exact test or χ2 test, a sample size of 45 in each group allowed detection of a ≥25% rate difference at an α level of .05, with statistical power of 72%.

This study protocol was approved by the human use committee at Tripler Army Medical Center. Investigators adhered to the policies for protection of human subjects prescribed in the Code of Federal Regulations, Title 45, Part 46.

RESULTS
One hundred twenty-seven patients were screened for eligibility. Of these, 7 were excluded because they were not able to meet the residency requirement of 1 year or their families were not interested in participating. Sixty were assigned randomly to the office-based group and 60 to the virtual group (Fig 1). Enrollment began in March 2003 and ended in December 2003. Patient data collection ended with the last patient’s final visit in February 2005.

Patients in the office-based group were slightly younger, and more had mild persistent asthma. There were more dropouts (defined as patients who were non-adherent/unable to be contacted for 8 weeks or families who notified investigators of an unanticipated military-related relocation) in the virtual group (Table 1).

Total study participation time exceeded 54 weeks for 10 patients in the virtual group and 17 patients in the office-based group, because of difficulty in scheduling the final visit. Many of the study families had a parent deploy to the war during the study year. Data were collected for all patients until the final visit was completed, so that the mean study time was 10.6 ± 3.3 months in the virtual group and 11.9 ± 2.1 month in the office-based group. Therefore, there were a total of 53 study years, 636 study months, or 19 329 study days in the virtual group; there were 59.4 study years, 713 study months, or 21 675 study days in the office-based group. The total study time was greater in the office-based group than in the virtual group (P < .05).

Both groups achieved excellent therapeutic adherence. Computer-accessed data on use of asthma controller medications (inhaled corticosteroids) were not different between groups, at 1 inhaler per month. Submission of inhaler technique videos was only one third of that expected from the research protocol, despite regular contact by case managers (Table 2).

Diagnostic adherence was variable. Although there was significantly greater symptom diary adherence in the virtual group (35.4%, compared with 20.8%; P < .01), no diary entry was recorded for 60% to 80% of study days. On the whole, patients completed symptom diary entries only every 2.8 days in the virtual group and 4.8 days in the office-based group. On the basis of these limited entries, patients in the virtual group recorded
61.1 ± 29.6 symptom-free days. Patients in the office-based group recorded 51.7 ± 37.6 symptom-free days. Electronic submission of peak flow measurements by patients in the virtual group was only one fourth of that directed by the protocol, despite regular contact by case managers. In both groups, peak flow technique and

### TABLE 1 Demographic Features

<table>
<thead>
<tr>
<th></th>
<th>Virtual Group</th>
<th>Office-Based Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>Male, n</td>
<td>37</td>
<td>38</td>
</tr>
<tr>
<td>Female, n</td>
<td>23</td>
<td>22</td>
</tr>
<tr>
<td>Age, mean ± SD, y</td>
<td>10.2 ± 3.1*</td>
<td>9.0 ± 3.0*</td>
</tr>
<tr>
<td>Mild persistent asthma, n</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>Moderate asthma, n</td>
<td>41</td>
<td>40</td>
</tr>
<tr>
<td>Severe asthma, n</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Dropouts, n</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Baseline FVC, mean ± SD, % predicted</td>
<td>103.7 ± 17.4</td>
<td>104.5 ± 15.4</td>
</tr>
<tr>
<td>Baseline FEV1, mean ± SD, % predicted</td>
<td>104.1 ± 19.9</td>
<td>96.8 ± 13.04</td>
</tr>
<tr>
<td>Baseline FEF25–75, mean ± SD, % predicted</td>
<td>83.8 ± 25.6</td>
<td>84.3 ± 23.5</td>
</tr>
</tbody>
</table>

FVC indicates forced vital capacity; FEV1, forced expiratory volume in 1 second; FEF25–75, forced expiratory flow in midexpiratory phase.

*P < .05.

### TABLE 2 Therapeutic Adherence Measures

<table>
<thead>
<tr>
<th></th>
<th>Virtual Group</th>
<th>Office-Based Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total No. of controller inhalers refilled</td>
<td>636</td>
<td>771</td>
</tr>
<tr>
<td>Participation, patient-mo</td>
<td>636*</td>
<td>713*</td>
</tr>
<tr>
<td>Controller inhaler use, inhalers per patient-mo</td>
<td>1.0</td>
<td>1.1</td>
</tr>
<tr>
<td>Total No. of inhaler videos sent</td>
<td>996</td>
<td>NA</td>
</tr>
<tr>
<td>Inhaler videos submitted, No. per patient-mo</td>
<td>1.56</td>
<td>NA</td>
</tr>
<tr>
<td>Expected inhaler videos, No. per mo</td>
<td>4.8</td>
<td>NA</td>
</tr>
<tr>
<td>e-check (inhaler video submission) adherence, %</td>
<td>33.1</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA indicates not applicable.

*P < .01.
percentage of personal best were consistently good (Tables 3 and 4).

Disease control was excellent in both groups. Emergency department visits and hospitalizations were rare. There were only 44 unscheduled, asthma-related, office visits in the virtual group (1.2 visits per patient-year) and 47 visits in the office-based group (1.3 visits per patient-year). There were no differences in rescue therapy use between groups. There were 309 β-receptor agonist refills in the virtual group (5.8 refills or canisters per patient-year) and 323 in the office-based group (5.4 refills or canisters per patient-year). There were 133 prednisone bursts in the virtual group (2.5 bursts per patient-year) and 129 prednisone bursts in the office-based group (2.2 bursts per patient-year). Mean peak flow measurements (as percentage of personal best) were similar for the 2 groups. There were no differences between groups in forced vital capacity, forced expiratory volume in 1 second, or forced expiratory flow in midexpiratory phase at the conclusion of the study (Table 4).

Asthma knowledge scores increased significantly in both groups (10.0 ± 10.2% increase in the virtual group and 9.2 ± 7.5% increase in the office-based group). There was no difference in asthma knowledge test improvement between the groups. Pediatric asthma quality of life scores measured for caregivers improved significantly during the course of the study for both groups, from 5.7 ± 1.1 to 6.4 ± 1.0 for parents of children in the virtual group and from 5.5 ± 1.1 to 6.2 ± 0.8 for parents of children in the office-based group ($P < .05$, 2-way analysis of variance). There was no difference in the quality of life score changes for children in the 2 groups. Quality of life scores for either children or parents did not differ between the office-based and virtual groups (Table 5).

Mean patient scores for DPI technique improved for both office-based and virtual groups from baseline to both 26 and 52 weeks. The mean score for the office-based group improved from 82.5 ± 16% to 94.3 ± 10.9% at 26 weeks and 92.8 ± 13.9% at 52 weeks ($P < .01$). The mean DPI technique score for the virtual group
improved from 86.7 ± 16.2% to 97 ± 9.7% at 26 weeks and 97 ± 9.2% at 52 weeks (P < .01). There was no significant change in DPI technique scores between 26 and 52 weeks for either group. There was no difference between DPI technique scores for the office-based group and the virtual group at baseline, 26 weeks, or 52 weeks (Fig 2).

Mean patient scores for MDI-VHC technique improved for the office-based group between baseline and 26 weeks (from 86 ± 16.6% to 93 ± 11.4%; P < .05) but returned to baseline at 52 weeks (from 86 ± 16.6% to 89 ± 15%; not significant). The mean MDI-VHC technique score for the virtual group did not change significantly between baseline and 26 weeks (87 ± 14.4% and 92 ± 12.2%, respectively; not significant) but was significantly greater at 52 weeks (94 ± 8%; P < .01). The mean MDI-VHC technique score was greater at 52 weeks for the virtual group than for the office-based group (94 ± 8% and 89 ± 15%, respectively; P < .05) (Fig 3).

DISCUSSION

Internet-based, in-home monitoring and patient asthma education were effective, compared with ideal office-based care, in this population of children of military service members. Both groups achieved excellent asthma control and had similar outcomes in what we think is the longest comparative trial of Internet-based, in-home patient monitoring and education to date.

Therapy for childhood asthma works. Patients who fail to respond to asthma therapy either do not have asthma or are not receiving the appropriate ambulatory therapy. For patients who have asthma uncomplicated by other patient factors (eg, gastroesophageal reflux) or accentuated by unhealthy living conditions (eg, constant exposure to environmental tobacco smoke), the failure to receive therapy is generally the result of nonadherence. In such cases, patients either do not receive or do not use the correct medication. Inadequate education of the patient, parent, and/or practitioner is inevitably a major component of the failure to receive appropriate therapy. This is especially true for the use of inhaled medications.

Adherence to the therapeutic plan was excellent for both groups, which likely was a major contributor to the outcomes. Pharmacy refill rates of inhaled corticosteroids reflected the use of 1 inhaler per month for patients in both the office and virtual groups, a rate consistent with that expected.

Hospitalization for treatment of asthma is the ultimate manifestation of a failure of ambulatory management for most children. At some level, it is a failure of the practitioner, patient, or parent to recognize and to respond to the child’s condition effectively. It may also represent a failure of the health care system.
ambulatory treatment of asthma should therefore be expected to result in a decrease in the number of asthma-related hospitalizations and rare use of unscheduled ambulatory care in the emergency department or clinic, as it did in patients from both study groups (1 hospitalization in each group).

In addition to refill rates for controller medication, therapeutic adherence included regular assessment of DPI and MDI-VHC techniques. High scores for both groups at intake suggested that the patients and their parents were well trained in asthma management before they began the study. We have emphasized asthma management at all primary care portals for our military population on Oahu for >10 years and have witnessed a decrease in the hospitalization rate of more than one half in the same period. It was not always so. In our previous research (the Telemedicine In-Home Monitoring Evaluation Project) conducted a decade earlier in the same population, only 30% of patients who were scored after viewing a teaching videotape had scores that reflected adequate MDI-VHC technique. These data suggest that asthma management and outcomes can be improved in a population of children over time.

Children in both the virtual and office-based group who used DPI corticosteroids to control their asthma had consistently good technique throughout the yearlong study, with no significant decrease in technique. However, children who used MDI-VHC to deliver corticosteroids experienced improvement in technique in the first 26 weeks of the study. The improvement was likely related to frequent monitoring and feedback, both virtually and in person for the office-based group. Children in the office-based group experienced a decrease in their technique during the second 26 weeks, when they had only weekly telephone contact unless they were seen for an unscheduled sick visit with their provider or case manager. Both the degree of increase in MDI-VHC scores and the statistical difference between the scores of the virtual and office-based groups at 52 weeks may be of limited clinical significance. However, the difference suggests the possibility that infrequent monitoring may lead to a decline in inhaler technique over time.

As little as 5% of the drug is delivered to children using MDI-VHC under ideal conditions. Therefore, a small decline in inhaler technique can lead to significantly diminished drug effectiveness. Poor MDI-VHC technique is a frequent reason for treatment failure in children. There may be a role for the use of telemedicine to monitor inhaler technique in children between regular office visits or even to improve monitoring of children separate from their primary care provider or pulmonary specialist.

Patients refilled an average of 1 β-receptor agonist inhaler every other month. This is a relatively high rate for rescue medication use for children whose asthma is well controlled, although accepted guidelines suggest that occasional daily use of rescue medication may be acceptable. It is not clear how the children used the β-receptor agonist inhalers, whether some were for school use during exercise, or even whether other family members used some of the inhalers.

Asthma action plans directed the use of orally administered corticosteroids under specific conditions of persistent symptoms, as well as the indications to seek medical therapy. Children in the virtual group had automated plans linked to their electronic, personalized, Web pages. Children in the office-based group had paper copy plans updated at each visit. There were similar numbers of oral corticosteroid rescue treatments for children in the 2 groups, which reflects the use of the asthma action plan and home corticosteroid therapy. There were also similar numbers of unscheduled asthma clinic visits and emergency department visits in the 2 groups, which reflects the equivalent clinical limitations of the asthma action plans in the 2 groups.

Children and their parents began the study with excellent knowledge of asthma. In both the virtual and office-based groups, the subjects’ asthma knowledge improved over the course of the year, which reflects the effectiveness of both electronic and in-person asthma education. Neither form of education proved superior, although it is difficult to generalize because of the high baseline asthma knowledge. The parents’ quality of life questionnaire scores improved for both groups, whereas the children’s scores did not. Perhaps the parents’ improved understanding of their child’s disease and their perception of improved asthma control and access to care contributed to the improved quality of life. Although virtual group patients were older than office-based group patients, we do not think that a 1-year age difference was a developmental factor in this population.

Despite the excellent outcomes for both groups, adherence to submission of MDI-VHC or DPI technique videos was disappointing. Only one third the number of anticipated videos were submitted to the Web site and scored. Peak flow meter video adherence was only one fourth of that expected. Similarly disappointing was the poor adherence in the use of the asthma diary by either the virtual group or the office-based group; although there were 50% more entries in the virtual asthma diaries, only one third of possible entries were completed. The relatively poor adherence with the diary explains the low number of symptom-free days in both groups; parents and children tended to remember to complete the diary during times when they were experiencing symptoms. The inconsistent adherence with the video submissions may reflect decreased interest in the Web site as the novelty waned. On average, use of the patient Web site declined steadily throughout the study period, and there was wide variation in use of the site between patients (Table 6). Poor adherence with the traditional and novel monitors of asthma management
used in the 2 groups under these research conditions raises concerns about their use for clinical decision-making under normal practice conditions. It is best not to depend on a single measure for clinical decision-making and, when in doubt, to contact or to examine the child.

This study was not designed for detailed cost analysis because the economics of practice within a military setting are not always comparable to those in the civilian sector. In addition, the office-based care was designed purposely as standardized and thus more expensive than usual office care. We attempted to design the best asthma care possible to provide the most stringent comparison with in-home monitoring. There were 337 scheduled and unscheduled visits in the office-based group and 157 in the virtual group. Even in this population of children with excellent asthma control and equal access to care, parents of office-based group patients spent nearly 20 hours/year away from work caring for their children with asthma, with round-trip travel time of ~1.5 hours for the visit and hospital encounter. Parents of virtual group patients spent one half of that.

The Asthma In-Home Monitoring project was based on the assumption that technology for in-home monitoring will become increasingly available in the decades to come. According to a Harris poll, 65% of the households in the United States are online and 44% have broadband Internet access. It is reasonable to expect that the majority of families will have Internet access in the future and that it will become a cost-effective way to monitor children with asthma, to avoid unscheduled emergency department visits and hospitalizations. Because it allows for daily monitoring of patients, there is an opportunity to increase adherence to treatment regimens, although the reliability of these measures should be subject to the same scrutiny and even skepticism that would be used for any of the traditional ambulatory measures used to monitor patients with asthma.

Internet-based telemedicine is an effective adjunctive method to monitor children with asthma at home. Increasing availability of low-cost technology and Internet access will very likely add this capability to the armamentarium of ambulatory management tools for children with asthma. It may be of special significance for families separated from specialty care by long distances. These tools should supplement and extend the relationships between patient, parent, and provider, which remain key to the successful management of chronic diseases in childhood.

ACKNOWLEDGMENTS

This research was supported by a grant from the US Army Medical Research Acquisition Activity.

We thank the late Francis Malone, MD, for his efforts in protocol and grant writing; Catherine Uyehara, PhD, for assistance with statistical analysis; Michael Darnall, Melcolm Goto, and Branden Tanga from the Pacific Telehealth and Technology Hui (Honolulu, HI) for technical assistance; and Tina Brads-Pitt, RN, for assistance with preliminary study tasks.

REFERENCES


### TABLE 6 Virtual Group Web Site Access Time After Initial Web Site Training

| Access Time, min | Weeks 0–2 | Weeks 2–6 | Weeks 6–12 | Weeks 12–26 | Weeks 26+
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>7230</td>
<td>4637</td>
<td>4736</td>
<td>6643</td>
<td>7195</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>121 ± 251</td>
<td>77 ± 91</td>
<td>79 ± 100</td>
<td>113 ± 144</td>
<td>156 ± 143</td>
</tr>
<tr>
<td>Time per patient per wk</td>
<td>60.5</td>
<td>19.3</td>
<td>13.2</td>
<td>8.1</td>
<td>6.0</td>
</tr>
</tbody>
</table>

Access Time, min
PLAN TO SHRINK HOSPITAL INDUSTRY HAS STATEWIDE EFFECT

“A long-awaited plan to shrink New York State’s hospital industry landed yesterday with force, with proposals that could effectively eliminate 20 or more hospitals and thousands of jobs, and force dozens of other hospitals to shrink, merge or take on new roles. The recommendations by the Commission on Health Care Facilities in the 21st Century go far beyond the nine hospital closings and downsizings that state officials reported late Monday, after being briefed by the commission. Several other hospitals would cease to exist through merges or conversion to new uses, and more could be eliminated if they refuse to merge. In addition, the commission reached deep into the particular operations of many individual institutions, ordering them to eliminate specific numbers of beds, telling some to eliminate psychiatric, substance abuse or maternity wards, and in some cases directing others to take on those functions. The changes would mark an epochal shift in an industry that, person for person, is much bigger in New York than in other states, and traditionally has fiercely resisted shrinkage. But hospitals are also in worse financial shape in New York than in any other state, having lost money for eight consecutive years.”

Noted by JFL, MD
OBJECTIVE. The goal was to determine the predictors of a prolonged course for children with acute otitis media.

METHODS. A meta-analysis of data with the observation groups of 6 randomized, controlled trials was performed. Participants were 824 children, 6 months to 12 years of age, with acute otitis media. The primary outcome was a prolonged course of acute otitis media, which was defined as fever and/or pain at 3 to 7 days.

RESULTS. Of the 824 included children, 303 had pain and/or fever at 3 to 7 days. Independent predictors of a prolonged course were age of <2 years and bilateral acute otitis media. The absolute risk of pain and/or fever at 3 to 7 days for children <2 years of age with bilateral acute otitis media (20% of all children) was 55%, and that for children ≥2 years of age with unilateral acute otitis media (47% of all children) was 25%.

CONCLUSIONS. The risk of a prolonged course was 2 times higher for children <2 years of age with bilateral acute otitis media than for children ≥2 years of age with unilateral acute otitis media. Clinicians can use these features (ie, age of <2 years and bilateral acute otitis media) to inform parents more explicitly about the expected course of their child’s otitis media and to explain which features should prompt parents to contact their clinician for reexamination of the child.
ACUTE OTITIS MEDIA (AOM) is one of the most common childhood infections, the leading cause of doctors’ visits, and the most frequent reason children receive antibiotics or undergo surgery. The high incidence of and high rate of spontaneous recovery from AOM suggest that it is a natural phenomenon, inevitable (like a common cold), and part of the gradual maturation of the child’s anatomy and immune system. However, untreated AOM can lead occasionally to supplicative complications, such as acute mastoiditis.

The treatment of AOM is still controversial. Many children are given antibiotics, although systematic reviews suggest that there is only marginal benefit for most children. An estimated 8 to 17 children need to be treated for 1 child to benefit from earlier resolution of symptoms. The effects of prescription of antibiotics are important, because prescription could increase antibiotic resistance, increase revisit rates, and increase the likelihood of seeking medical care for future illnesses.

Currently we have no tools to discriminate between children with mild, self-limiting episodes of AOM and those at risk of a prolonged course. To date, only Little et al have examined which children with AOM are at risk for a prolonged course. They showed that the presence of fever and vomiting increased the risk of a prolonged course. Prognostic studies require relatively large numbers of participants developing the outcome of interest (in general, 10 participants per prognostic determinant studied). The power of the 2 earlier studies was therefore very limited. We determined the predictors of a prolonged course from the combined individual patient (control) data of 6 randomized, controlled trials.

METHODS
Selection of Trials
A systematic literature search was performed with PubMed, Embase, the proceedings of International Symposia on Recent Advances in Otitis Media, and the Cochrane Library. To be selected for the individual patient data meta-analysis, trials needed to be randomized, they needed to include children 6 months to 12 years of age with AOM, and the comparison needed to be between antibiotics and a placebo or no treatment (observation group). The primary investigators of all selected trials were asked for the raw data from their trials.

Patients
Only patients in the observation groups of the available randomized, controlled trials were selected for this prognostic study.

Outcomes
The primary outcome was a prolonged course of AOM, which was defined as pain and/or fever at 3 to 7 days. We used this composite end point because both factors are relevant from clinical and patient (parental) perspectives. Fever was defined as a temperature of ≥38°C, and pain (yes versus no) was measured with diaries completed by the parents. Both outcome measures needed to be dichotomized, because several trials measured them only in that way. Fever and pain were also studied separately (secondary outcomes).

Predictors
On the basis of a literature search and the availability of information in routine clinical practice, the following baseline candidate predictors were selected: age (<2 years versus ≥2 years), gender (boys versus girls), season (autumn/winter versus spring/summer), having been breastfed (yes versus no), smoking in the household (yes versus no), siblings (yes versus no), family history of AOM (yes versus no), recurrent AOM (yes versus no), fever (yes versus no), pain (yes versus no), bilateral AOM (yes versus no), otorrhea (yes versus no), runny nose (yes versus no), crying (yes versus no), coughing (yes versus no), red tympanic membrane (yes versus no), bulging tympanic membrane (yes versus no), and perforation of the tympanic membrane (yes versus no).

Statistical Methods
To determine whether pooling was justified, heterogeneity between studies was assessed with the $I^2$ statistic. Because the $P$ value was <25%, pooling was performed.

The association between each prognostic factor and the presence or absence of fever and/or pain at 3 to 7 days was examined with univariate logistic regression analyses. Predictors that were associated with the outcome in univariate analyses ($P ≤ .10$) were included in multivariate logistic regression analyses. The model was reduced through exclusion of predictors with $P$ values of >.05. The predictive accuracy of the models was estimated on the basis of their reliability (goodness of fit) with Hosmer-Lemeshow tests. The model’s ability to discriminate between children with and without poor outcomes was estimated as the area under the receiver operating characteristic curve of the model. The receiver operating characteristic curve area is a suitable parameter to summarize the discriminative or predictive value and can range from 0.5 (no discrimination, like a coin flip) to 1.0 (perfect discrimination). In addition, we calculated the absolute risks of a prolonged course across combinations of independent predictors.

We also calculated the proportions of children with a prolonged course on each consecutive day within each subgroup of the identified independent predictors. The resulting survival curves were based on diary records of the presence of symptoms completed by the parents in 5 of the 6 included trials. Finally, sensitivity analyses, including only those trials that measured the outcomes
on the same day, used the same dosage, or included a placebo, were performed.

**Missing Values**

Information was available for 72% of the predictor variables (range: 28%–100%) and for 90% of the outcome variables (range: 81%–98%). Data are seldom missing at random; it has been shown that removal of subjects with a missing value for ≥1 of the predictors studied (complete case analysis) commonly leads to biased results and certainly to loss of power.\(^{14,15}\) To decrease bias and to increase statistical efficiency, it is better to impute missing data than to perform a complete case analysis. Accordingly, we imputed the missing data for each trial by using the linear regression method (missing value analysis) available in SPSS 12.0 for Windows (SPSS, Chicago, IL). Such imputation is based on the correlation between each variable with missing values and all other variables, as estimated for the set of complete subjects. We imputed missing values only within trials. Some predictor and outcome variables are therefore still missing because they were not measured at all in 1 of the included trials.

**RESULTS**

We identified 19 trials that studied the effectiveness of antibiotics in children with AOM. Nine trials were excluded because of inadequate randomization or lack of availability of information on the outcomes included in our meta-analysis or because the trial was conducted with a special study population.\(^{16–24}\) The principal investigators of 6 trials provided their data.\(^{25–30}\) The data from the other 4 trials were not available.\(^{31–34}\) The mean number of children studied in the 6 included trials ranged from 121 to 512. The mean age was 3.4 years (range: 6 months to 12 years); 50% of the children were boys, 52% had recurrent AOM, and 27% had bilateral AOM.

In total, 824 patients (ie, all children with AOM who were assigned to the observation group in the 6 available randomized trials) were included for determination of the predictors of a prolonged course. The characteristics of these 824 included children are shown in Table 1. Of the 824 included children, 91 (11%) had fever at 3 to 7 days, 257 (31%) had pain at 3 to 7 days, and 303 (37%) had pain and/or fever at 3 to 7 days.

Table 2 shows the univariate predictors of both the primary and secondary outcomes. Univariate predictors of pain and/or fever at 3 to 7 days were age of <2 years, family history of otitis media, fever at inclusion, bilateral AOM, otorrhea, and red tympanic membrane. Predictors of fever at 3 to 7 days were age of <2 years, winter season, fever at inclusion, pain at inclusion, bilateral AOM, runny nose, and abnormal tympanic membrane. Predictors of pain at 3 to 7 days were age of <2 years, having siblings, bilateral AOM, otorrhea, and red tympanic membrane.

Table 3 shows the independent predictors of pain and/or fever, pain, and fever at 3 to 7 days. Independent predictors of pain and/or fever at 3 to 7 days were age of <2 years (odds ratio [OR]: 2.1; 95% confidence interval [CI]: 1.5–2.9) and bilateral AOM (OR: 1.7; 95% CI: 1.2–2.4). The prognostic model showed a good fit (goodness-of-fit test, \(P = .93\)), and the AUC was 0.63 (95% CI: 0.59–0.68). Independent predictors of pain at 3 to 7 days were age of <2 years (OR: 2.0; 95% CI: 1.4–2.9) and bilateral AOM at baseline (OR: 1.8; 95% CI: 1.3–2.6). The goodness of fit of this model was good (\(P = .59\)), and the AUC was 0.64 (95% CI: 0.59–0.68). Age of <2 years (OR: 1.6; 95% CI: 1.0–2.6) and fever at baseline (OR: 3.0; 95% CI: 1.8–4.9) were independent predictors of fever at 3 to 7 days. The goodness-of-fit test for this model indicated a good fit of the prognostic model (\(P = .92\)), and the AUC was 0.67 (95% CI: 0.61–0.73). There was no significant interaction between the independent predictors (all \(P > .6\)).

---

**TABLE 1**

**Characteristics of the 824 Included Children**

<table>
<thead>
<tr>
<th>Personal factors</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;2 y</td>
<td>287 (35)</td>
</tr>
<tr>
<td>≥2 y</td>
<td>537 (65)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>411 (50)</td>
</tr>
<tr>
<td>Girls</td>
<td>413 (50)</td>
</tr>
<tr>
<td><strong>Season</strong></td>
<td></td>
</tr>
<tr>
<td>Autumn/winter</td>
<td>620 (75)</td>
</tr>
<tr>
<td>Spring/summer</td>
<td>204 (25)</td>
</tr>
<tr>
<td><strong>Being breastfed</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>255 (31)</td>
</tr>
<tr>
<td>No</td>
<td>143 (17)</td>
</tr>
<tr>
<td><strong>Recurrent AOM</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>429 (52)</td>
</tr>
<tr>
<td>No</td>
<td>395 (48)</td>
</tr>
<tr>
<td><strong>Symptoms at baseline</strong></td>
<td></td>
</tr>
<tr>
<td>Fever at inclusion</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>287 (35)</td>
</tr>
<tr>
<td>No</td>
<td>419 (51)</td>
</tr>
<tr>
<td>Pain at inclusion</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>724 (88)</td>
</tr>
<tr>
<td>No</td>
<td>100 (12)</td>
</tr>
<tr>
<td><strong>Bilateral AOM</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>220 (27)</td>
</tr>
<tr>
<td>No</td>
<td>440 (53)</td>
</tr>
<tr>
<td><strong>Otorrhea</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>65 (8)</td>
</tr>
<tr>
<td>No</td>
<td>268 (27)</td>
</tr>
<tr>
<td><strong>Tympanic membrane</strong></td>
<td></td>
</tr>
<tr>
<td>Red</td>
<td>754 (91)</td>
</tr>
<tr>
<td>No</td>
<td>70 (9)</td>
</tr>
<tr>
<td>Bulging</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>342 (42)</td>
</tr>
<tr>
<td>No</td>
<td>482 (58)</td>
</tr>
<tr>
<td>Perforated</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>19 (2)</td>
</tr>
<tr>
<td>No</td>
<td>268 (33)</td>
</tr>
</tbody>
</table>

The percentages do not always add to 100 because of missing data.
Figure 1 shows the proportions of children experiencing fever and/or pain during the follow-up period. Table 4 shows the absolute risks of fever and/or pain, fever, and pain at 3 to 7 days for children with certain combinations of independent prognostic factors. The absolute risk of pain and/or fever at 3 to 7 days was highest for children <2 years of age with bilateral AOM (ie, 55%; 95% CI: 47%–63%; 20% of all children). The risk for children ≥2 years of age with unilateral AOM was 25%; 95% CI: 20%–30%; 47% of all children). The results of the complete case and sensitivity analyses (including only the trials that measured the outcome at the same time, used the same dosage, or included a placebo) were in agreement with the overall results.

### DISCUSSION

Combining data from the observation groups of 6 randomized trials, we found that age of <2 years and bilateral AOM were independent predictors of a prolonged course of AOM. The absolute risk of pain and/or fever at 3 to 7 days for children <2 years of age with bilateral AOM (20% of all children) was 55%, whereas the risk for children ≥2 years of age with unilateral AOM (47% of all children) was 25%.

Reliably identifying prognostic factors for a prolonged course has been difficult, because individual studies were too small for valid reliable analyses. Meta-analyses of the original data of the individual trials have been proposed as a major improvement, because they have...
greater power for informative prognostic analyses, allowing more-thorough assessment of whether differences are spurious.\textsuperscript{35} By reanalyzing the data of 6 trials, we were able to include 824 children. Moreover, with exclusion of the children from the antibiotic treatment groups, the results represent the natural history of untreated AOM. To our knowledge, this is the first study that attempts to predict the absolute risks of a prolonged course in children with AOM.

Some of our findings deserve additional discussion. First, 6 of the 10 relevant randomized trials were included in our meta-analysis. The main-effect results for the 4 trials whose individual patient data were not available were very similar to those for the included trials (ie, antibiotics had a marginal effect); therefore, it is not expected that inclusion of these data would have changed the results of this meta-analysis. The major advantage of such a large study is that it minimizes the possibility of both type I and II errors. Because our study included 300 children who developed the outcome, we had the power to evaluate ~30 predicting variables, whereas we studied only 18.

Second, the results are based on data for children participating in trials, who may not be representative of those visiting general practitioners. For example, more severely affected children may be underrepresented in the trials. However, because we had access to raw data from 6 randomized trials, the numbers in the specific high-risk groups that are often underrepresented in a single trial were higher than in all studies performed to date. Furthermore, the population of children included in our meta-analysis seemed representative of those visiting general practitioners in daily life, because the proportions of children <2 years and ≥2 years of age in our study were similar to those in national surveys (ie, 35% vs 33% and 65% vs 67%, respectively).\textsuperscript{36}

Third, because not all trials used objective diagnostic methods (eg, pneumatic otoscopy or tympanometry), some children might not have suffered from an ear infection. However, results of sensitivity analyses with the 3 trials that did use these diagnostic methods\textsuperscript{27,29,30} were in agreement with the overall results.

Fourth, we included only children allocated to the observation arm in our analyses, which is appropriate because treatment with antibiotics would influence the course of the disease and result in an invalid natural history model. Fifth, some predictor and outcome variables (eg, fever and pain) might have been more informative if analyzed on a continuous scale. Some trials did measure these items on a continuous scale but, because others did not, we needed to recode these items as dichotomous variables. Other possible important outcomes, such as severity of pain, night disturbance, and

<table>
<thead>
<tr>
<th>Predicting Variables</th>
<th>No. (%) of All Children</th>
<th>Absolute Risk (95% CI), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain and/or fever at 3–7 d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of &lt;2 y plus bilateral AOM</td>
<td>134 (20)</td>
<td>55 (47–63)</td>
</tr>
<tr>
<td>Age of &lt;2 y plus unilateral AOM</td>
<td>132 (20)</td>
<td>40 (32–48)</td>
</tr>
<tr>
<td>Age of ≥2 y plus bilateral AOM</td>
<td>86 (13)</td>
<td>35 (25–45)</td>
</tr>
<tr>
<td>Age of ≥2 y plus unilateral AOM</td>
<td>308 (47)</td>
<td>25 (20–30)</td>
</tr>
<tr>
<td>Pain at 3–7 d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of &lt;2 y plus bilateral AOM</td>
<td>134 (20)</td>
<td>46 (38–54)</td>
</tr>
<tr>
<td>Age of &lt;2 y plus unilateral AOM</td>
<td>132 (20)</td>
<td>33 (25–41)</td>
</tr>
<tr>
<td>Age of ≥2 y plus bilateral AOM</td>
<td>86 (13)</td>
<td>30 (20–40)</td>
</tr>
<tr>
<td>Age of ≥2 y plus unilateral AOM</td>
<td>308 (47)</td>
<td>19 (15–23)</td>
</tr>
<tr>
<td>Fever at 3–7 d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of &lt;2 y plus fever</td>
<td>153 (22)</td>
<td>23 (16–30)</td>
</tr>
<tr>
<td>Age of &lt;2 y plus no fever</td>
<td>134 (19)</td>
<td>8 (3–13)</td>
</tr>
<tr>
<td>Age of ≥2 y plus fever</td>
<td>134 (19)</td>
<td>15 (9–21)</td>
</tr>
<tr>
<td>Age of ≥2 y plus no fever</td>
<td>285 (40)</td>
<td>6 (3–9)</td>
</tr>
</tbody>
</table>

Percentages may change because of missing data.
duration of symptoms before study entry, were not measured in all trials and therefore could not be studied. Consequently, we cannot rule out the possibility that some children had relatively mild complaints.

To date, 2 studies that tried to identify predictors of antibiotic use have been performed.\textsuperscript{5,37} They also found that fever was a prognostic factor for antibiotic usage (indicating a poor outcome). Neither of them, however, studied bilateral AOM, and Little et al\textsuperscript{9} dichotomized age at 3 years, which seemed not to predict a poor outcome. A cutoff value of 2 years of age seems to be more relevant, because it is known that protective immunity against infections with encapsulated bacteria, such as the species that cause AOM, depends on the ability to produce specific antibodies against bacterial capsular polysaccharides, which is inadequate until 2 years of age.\textsuperscript{18}

The anatomic features of the eustachian tubes and the nasopharynx also differ with age.\textsuperscript{39,40} Consequently, children <2 years of age seem to be more susceptible to AOM. Moreover, guidelines often recommend treating children <2 years of age with antibiotics, whereas an initial observation period is recommended for children ≥2 years of age.\textsuperscript{41–43} We also studied different cutoff points, which showed that children <3 years of age also had a higher risk of a prolonged course but the absolute risks were much smaller (ie, the absolute risk of pain and/or fever at 3–7 days for children <3 years of age with bilateral AOM was 33%).

CONCLUSIONS

The risk of fever and/or pain at 3 to 7 days was 2 times as high for children <2 years of age with bilateral AOM, compared with children ≥2 years of age with unilateral AOM. Clinicians can use these features (ie, age of <2 years and bilateral AOM) to inform parents more explicitly about the expected course of their child’s AOM and to explain which features should prompt parents to contact their clinician for reexamination of the child.

ACKNOWLEDGMENTS

The study was funded by the Dutch College of General Practitioners and the Netherlands Organization for Health Research and Development (grant 4200.0010). We thank Anne Schilder for commenting on an earlier version of this article.

REFERENCES

19. Rudberg RD. Acute otitis media: comparative therapeutic results of sulphonamide and penicillin administered in various forms. \textit{Acta Otolaryngol.} 1954;113(suppl):1–79
COMMENTARY

Pain Assessment in Preterm Neonates

K. J. S. Anand, MBBS, DPhil

Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock, Arkansas

The author has indicated he has no financial relationships relevant to this article to disclose.

Accurate pain assessment is a key and central issue that confronts clinicians at the bedside of preterm neonates or researchers who study nonverbal subjects.1 Although many validated methods for pain assessment are available,2 none of them are widely accepted or clearly superior to others. Consequently, no “gold standard” can be recommended for broad adoption in clinical practice because of 2 problems that are common to all assessment methods.

First, these methods were developed from studies of neonates who underwent acute painful procedures (heal stick, venipuncture, circumcision). Physiologic or behavioral parameters chosen for inclusion in these methods were those that changed most acutely in response to tissue injury and subsided after painful stimulation was over. These responses were thought to be “specific” for neonatal pain. Subsequent research, however, noted that many infants do not produce “specific” responses when exposed to invasive, skin-breaking procedures.3 Preterm newborns who were more immature, asleep, or exposed to previous painful procedures were less likely to demonstrate specific responses to pain,3 whereas previous physical handling accentuated their responses to acute pain.4

Second, these behavioral and physiologic responses require the subjective evaluation of a clinical observer.5 Significant interobserver variability occurs6,7 and can be reduced but not eliminated by training or greater experience. The need for training creates a significant obstacle for the routine use of these methods. Because physicians or nurses who rotate temporarily through the unit are not trained, most neonates are not assessed for pain, which greatly reduces their likelihood of receiving analgesia.8,9

Most pain measures incorporate specific facial movements associated with pain expression in neonates; comparative studies have noted greater accuracy for these versus other pain parameters.10 Brahnam et al11 have developed a neural network–based learning algorithm trained on a database of 204 photographs, of which 60 were obtained from infants in pain. This system, called the Classification of Pain Expressions (COPE), uses facial-recognition techniques to extract and examine features of an infant’s facial expression and was noted to have >90% accuracy. These authors are now working on a follow-up study involving 500 infants and using video images. Moving images should allow these researchers to investigate the dynamic characteristics of pain expressions, affording much greater accuracy and clinical utility.

Despite recent progress in assessing and treating acute pain in infancy, attempts to measure persistent or prolonged pain have been largely unsuccessful. Such infants remain untreated and may experience prolonged suffering, because many of the signs of acute pain are absent. The only tools that measure prolonged pain in infants include the N-PASS (Neonatal Pain, Agitation and Sedation Scale)12 and EDIN (Échelle Douleur Inconfort Nouveau-Né [Neonatal Pain and Discomfort Scale])13 (both of which have unproven construct validity for pain) and the DEGR (Douleur Enfant Gustave Roussy)14 (which has not been tested in preterm neonates). Other than the vital signs, physiologic parameters that represent autonomic activation were not included in pain-assessment methods, perhaps because they lack specificity for pain or because they require specialized equipment at the...
bedside: heart rate variability\textsuperscript{15} or sympathetic tone from spectral analyses of heart rate variability,\textsuperscript{16} changes in palmar sweating\textsuperscript{17} or skin conductivity,\textsuperscript{18} and changes in skin blood flow.\textsuperscript{19} Other methods such as pulsilometry\textsuperscript{20} have not been developed for neonates, whereas measuring salivary cortisol\textsuperscript{21} or other stress hormone responses\textsuperscript{22} are not practical for clinical use.

In addition, measures of sensory function were developed by using the tactile stimulation produced by Von Frey filaments,\textsuperscript{23,25} but these may only reflect hypersensitivity of the pain pathways and, at best, are indirect measures of pain during movement, dressing change, or physical examination. Reflex withdrawal after Von Frey stimulation does not lead to cortical activation either.\textsuperscript{26}

The mismatch between measured parameters for acute versus prolonged pain may be explained by 2 sequential phases in the neonate’s behavioral/physiologic responses to pain.\textsuperscript{27} These 2 phases are remarkably distinct, showing a psychophysiological “activation” or “shutdown,” but occur in succession, consistent with the sequential phases of “protest” and “despair” that have been described in studies of maternal separation.\textsuperscript{28,29} Because of their limited energy reserves, preterm infants cannot maintain the psychophysiological activation triggered by skin-breaking procedures if the pain becomes persistent.\textsuperscript{27} Consequently, methods of assessment developed from models of acute pain may not apply to patients with prolonged pain.

These difficulties may well be overcome in the near future. From evaluating ventilated preterm neonates who were randomly assigned to receive masked morphine or placebo infusions, Boyle et al\textsuperscript{10} found that facial expressions of pain, high activity levels, poor responses to handling, and lack of ventilator synchrony synchronically specifically identified neonates in the placebo group. These indicators, in addition to those listed above, may provide the initial parameters for specifically measuring persistent pain in neonates. Indicators for morphine analgesia, however, were not identified because clinicians found it easier to identify neonatal discomfort rather than analgesia.\textsuperscript{30}

Although neuroimaging techniques may partially define the supraspinal processing of pain, their subjective importance still has to be inferred.\textsuperscript{31} Bartocci et al\textsuperscript{12} used near-infrared spectroscopy to study cortical pain processing in 40 preterm neonates (born at 28–36 weeks of gestation) at 25 to 42 hours of age. Cortical activation was recorded over both somatosensory areas in 29 newborns and over the contralateral somatosensory and occipital areas in 11 newborns. Blood flow to somatosensory areas increased after standardized tactile (skin disinfection) and painful (venipuncture) stimuli but not to the occipital cortex, which implied a functional specificity for this response. Pain-related blood flow increases were more pronounced in male neonates, correlated with increasing postnatal ages and decreasing gestational ages, and appeared more prominent in the left (dominant?) somatosensory cortex.

Slater et al\textsuperscript{26} also recorded cortical activation after heel sticks in 18 infants between 25 and 45 weeks' gestation. Some neonates were studied repeatedly (1–5 times) between postnatal ages ranging from 5 to 134 days. Robust activation occurred over the contralateral somatosensory cortex, greater in awake than in asleep infants. No cortical response occurred after tactile stimulation of the heel even when accompanied by reflex limb withdrawal.\textsuperscript{26} Because neonatal responses to acute pain change with increasing exposure to invasive procedures after birth,\textsuperscript{33,35} these findings may reflect the effects of age-dependent cortical maturation or the cumulative experience of previous invasive procedures.

Taken together, these studies demonstrate robust cortical activation after acute pain in neonates, altered by gender, laterality, gestational age, postnatal age, behavioral state, previous pain experiences, and differences between reflex withdrawal and “pain.”\textsuperscript{26,32} The magnitude of this cortical response was reduced in 2 neonates who received morphine,\textsuperscript{26} which is the opposite of what may have been expected from their pain-assessment scores.\textsuperscript{36}

But, how can measured neurophysiological events represent a subjective experience? This is the hard problem,\textsuperscript{31,37} almost insurmountable in fetuses or preterm neonates.\textsuperscript{38} Yet, these data are important for 3 reasons. First, these nuanced activation responses at the highest level of sensory processing imply that preterm neonates may “consciously” perceive the acute pain of skin-breaking procedures. Second, similar neuroimaging methods or electroencephalography may provide ways of validating the currently available or novel pain-assessment methods that can be easily, reliably, routinely applied at the bedside. Third, newer methods designed to assess prolonged pain must adopt a different paradigm from those developed for invasive procedures. Until then, clinicians will need to estimate the pain that infants experience during neonatal intensive care and attempt to reduce it without altering their respiratory/hemodynamic stability or future brain development.

ACKNOWLEDGMENTS

I gratefully acknowledge financial support from the National Institute of Child Health and Human Development (grant HD50009), the National Center for Research Resources (grants RR018765 and RR016460), and the Arkansas Children’s Hospital Foundation.

REFERENCES

AlTHOUGH THERE is a growing body of literature criticizing the use of mere statistical significance as a measure of clinical impact, we submit that this concept has not been widely incorporated in the pediatric literature. This is especially problematic because an understanding of the limitations of using only statistical significance to evaluate treatments is necessary for readers of Pediatrics to draw accurate conclusions from data presented in this journal. Here we highlight some of the issues related to the complex problem of evaluating treatment effects and the importance of using clinical significance in addition to the traditional $P$ value.

Currently, the magical boundary of $P < .05$ holds great importance in whether a manuscript is accepted for publication, a research application is funded, or a new drug is approved by the Food and Drug Administration. We submit that if a treatment is to be useful to our children, it is not enough for treatment effects to be statistically significant; they also need to be large enough to be clinically meaningful. Evaluating treatment outcomes on the basis of $P$ value alone is problematic for several reasons. First, with a large sample, it is quite possible to have a statistically significant result between groups despite a minimal effect of treatment (ie, small effect size). Second, study outcomes with lower $P$ values are typically misinterpreted by pediatricians as having stronger effects than those with higher $P$ values. That is, most clinicians believe that a result with $P = .002$ has a much greater treatment effect than a result of $P = .045$. Although this is true if the sample size is the same in both studies, it is not true if the sample size is larger in the study with the smaller $P$ value. This confusion becomes particularly concerning when one realizes that most pharmacologically funded studies have very large sample sizes.

To combat overreliance on the $P$ value, we recommend that pediatricians be interested in answering 3 basic questions when examining the report of a clinical trial:

1. Could the findings of the clinical trial be solely a result of a chance occurrence? (ie, statistical significance)
2. How large is the difference between the primary end points of the study groups? (ie, impact of treatment, effect size)
3. Is the difference of primary end points between groups meaningful to a patient? (ie, clinical significance)

UNDERSTANDING STATISTICAL SIGNIFICANCE

As is familiar to most readers of Pediatrics, the $P$ value is the most commonly used method of evaluating the statistical significance of any finding. The origin of the $P$ value lies in 1925 with Sir Ronald A. Fisher, who first suggested the use of a boundary between significance and nonsignificance that was based on probability.1–3 He arbitrarily set this boundary at $P = .05$, where “$P$” stands for the probability that a finding of interest was reached by chance.1–2 Although Fisher’s emphasis on significance testing and the arbitrary boundary of $P < .05$ is familiar and widely used, it is important for pediatricians to recognize that this definition has been widely criticized.
over the past 80 years. Specifically, this approach is criticized because it does not take into account the size and clinical significance of the observed effect. That is, a small effect in a study with large sample size may have the same $P$ value as a large effect in a study with a small sample size.

In an attempt to address some of the limitations of the $P$ value, use of confidence intervals (CIs) has been advocated by some clinicians. It is important the readers realize, however, that these 2 definitions of statistical significance are essentially reciprocal. That is, a $P$ value of <.05 is essentially the same as having a 95% CI that does not overlap 0. CIs do have some advantage, however, in that they can be used to estimate the size of difference between groups. Unfortunately, this approach is not widely used in the pediatric literature, and CIs are mostly used today as surrogates for the hypothesis test rather than considering the full range of likely effect size.

**BEYOND THE $P$ VALUE: EFFECT SIZES**

Providing more information than either $P$ values or CIs, the group of statistics called “effect sizes” are measures of the magnitude of difference between groups, standardized by controlling for variation within groups. In other words, whereas a $P$ value denotes whether the difference between 2 groups in a particular study is likely to occur solely by chance, the effect size quantifies the amount of difference between these 2 groups. Because effect size is based on standardized differences between groups and not sample size, they better evaluate the strength of the intervention. Of particular relevance to pediatricians is effect sizes of the $d$ type, because these are primarily used to compare 2 treatment groups. $d$-type effect size is defined as the magnitude of difference between 2 means, divided by the SD ([mean of control group − mean of treatment group]/SD of the control group]. Thus, the $d$ effect size depends on variation within the control group and the differences between the control and intervention groups. Conventionally, $d$-type effect sizes that are near .20 are interpreted as small, effect sizes near .50 are considered “medium,” and effect sizes around .80 are considered “large.” Effect sizes of another type, the risk potency type, include likelihood ratios such as odds ratio, risk ratio, risk difference, and relative risk reduction. Clinicians are probably more familiar with these less abstract statistics, and it may be helpful to realize that likelihood statistics are a type of effect size. There are a number of different types of effect sizes, but description of these various types and formulae is beyond the scope of this commentary; however, we refer the interested reader to a number of review articles that discuss these issues.

**FURTHER STILL: CLINICAL SIGNIFICANCE**

At this point, we feel that it is important to caution *Pediatrics* readers that magnitude of change (effect size) should not be interpreted as an indication of clinical significance. The clinical significance of a treatment should instead be based on external standards provided by patients and clinicians. That is, a small effect size may still be clinically significant and, likewise, a large effect size may not be clinically significant. Indeed, there is a growing recognition that traditional methods used, such as statistical significance tests and effect sizes, should be supplemented with methods for determining clinical significant changes.

Although there is little consensus about the criteria for these efficacy standards, the most common definitions of clinically significant change include: (1) treated patients make a statistically reliable improvement in the change scores; (2) treated patients are empirically indistinguishable from a normal population after treatment; or (3) there are changes of at least 1 SD. The most frequently used method for evaluating the reliability of change scores is the Jacobson-Truax method in combination with clinical cutoff points. Using this method, change is considered unlikely to be the product of measurement error if the reliable change index is >1.96. That is, when the score of a patient has a change score >1.96, one can reasonably assume that indeed the score has improved.

The validity of each of the above-described methods can be improved further by establishing their external validity (eg, patient perspective). For example, Flor et al conducted a large meta-analysis and evaluated the effectiveness of multidisciplinary treatment for chronic pain. The investigators found that pain among the patients who received the intervention was reduced by 25% with an effect size of .7. Although this finding seems promising statistically, the meaning of the results change in light of findings from Colvin et al, who reported that patients consider only a 50% pain improvement a “treatment success.” Thus, in this example, a reduction of 25% in pain scores may be statistically but not clinically significant. Clearly, this is a developing area that warrants additional discussion.

**CONCLUSIONS**

The issue of clinical significance is of utmost importance to both pediatric researchers and clinicians. On the research side, it is imperative that studies routinely evaluate both statistical and clinical significance to advance our understanding of treatment effects. As such, we encourage researchers to report effect sizes, at the very least, and incorporate external validations of clinical significance when possible. On the clinical side, pediatricians must understand the potential disconnect between statistical and clinical significance when making decisions about the adoption of new treatments. The inter-
pretation of any research findings should occur in the context of the magnitude of change that occurred and the clinical significance of the findings.

ACKNOWLEDGMENTS
This work was supported in part by the National Institutes of Health through National Institute of Child Health and Human Development grant R01HD37007-02.

REFERENCES

TV DRUG ADVERTISING

“Two independent government watchdog groups sharply criticized consumer drug advertising recently, and a separate survey Jan 9 commissioned by the PricewaterhouseCoopers accounting and consulting firm indicated that skepticism is widespread among the public, too. Only 1 in 10 consumers said the direct-to-consumer, or DTC, ads could provide useful information to a large audience, the survey said. (Consumer drug advertising is not permitted in most of the world, except New Zealand and the United States.) The pharmaceutical industry itself acknowledges having an image problem. ‘It would be naïve to not acknowledge the fact that DTC advertising is also a lightening-rod in the health care debate in this country,’ said Billy Tauzin, the former congressman who is now president and chief executive of the Pharmaceutical Research and Manufacturers of America, in a speech to venture capitalists last spring. There is ‘one great problem’ that the manufacturers face,” he said: ‘in a word, it is trust.’ ‘While individual patients find the information useful in discussions with their physicians,’ he added in his speech, ‘patients, physicians and consumers generally express unhappiness with DTC advertising.’ . . . FDA officials said they had to deal with 54 000 drug promotions each year, aimed at both doctors and consumers.”

Noted by JFL, MD
Oxygen Therapy for Bronchiolitis

Joel L. Bass, MD*, David Gozal, MD

*Department of Pediatrics, Newton-Wellesley Hospital, Massachusetts General Hospital for Children, Harvard Medical School, Boston, Massachusetts; **Department of Pediatrics, Kosair Children’s Hospital Research Institute, University of Louisville, Louisville, Kentucky

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

A N AMERICAN ACADEMY of Pediatrics clinical practice guideline on the diagnosis and management of bronchiolitis was published recently in Pediatrics.1 This important guideline was evidence based and intended to help guide practitioners in the management of this commonly encountered and potentially serious condition.

Recommendation 7a in the guideline includes the following statement: “If the SpO2 [pulse oxygen saturation] does persistently fall below 90%, adequate supplemental oxygen should be used to maintain SpO2 at or above 90%. Oxygen may be discontinued if SpO2 is at or above 90%.”1

The authors state in their discussion of the recommendation that healthy infants have an SpO2 of >95% in room air, but they justify the lower cutoff for therapy on the grounds that, on the basis of the oxyhemoglobin dissociation curve, otherwise healthy infants gain little benefit in PaO2 with supplemental oxygen at SpO2 levels ≥90%. The subcommittee also suggested that clinicians consider maintaining a higher SpO2 in children with risk factors such as fever, acidosis, or some hemoglobinopathies. In the evidence profile that they cite as the basis for this recommendation, the benefit of use is stated as “shorter hospitalization,” and the potential harm is stated as “inadequate oxygenation.” Using these parameters, it is understandable how the subcommittee arrived at its recommendation.

It is unfortunate that the recommendation fails to address another significant consideration, viz, the impact of chronic or intermittent hypoxia on later cognitive and behavioral outcomes. A recently published review of the evidence on this topic2 documented a very convincing association between hypoxia and adverse cognitive outcomes in a wide variety of clinical circumstances including both short-term and intermittent exposures at SpO2 levels in the range of 90% to 94%. Although none of the articles cited in the review were specific to bronchiolitis, this was because there were no published studies relating to bronchiolitis, not because it has been studied and shown to be uniquely without consequence.

Given the overwhelming evidence in so many other clinical circumstances, we think it would have been prudent for the subcommittee to state that in other clinical situations adverse cognitive and behavioral outcomes have been reported at SpO2 levels between 90% and 94% and that this information should be taken into consideration when making decisions regarding administration of supplemental oxygen. Although this is not as clear cut as the 90% level, it is a more accurate representation of the potential risks of withholding supplemental oxygen therapy. This is an area in which physician discretion and informed parental input need to be considered and taken into account when making clinical decisions. Additional research into this issue should also be encouraged so that more definitive answers might be available in the future.

REFERENCES

Abbreviation: SpO2, pulse oxygen saturation

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.2006-3002
doi:10.1542/peds.2006-3002

Accepted for publication Oct 18, 2006

Address correspondence to Joel L. Bass, MD, Department of Pediatrics, Newton-Wellesley Hospital, 204 Washington St, Boston, MA 02462. E-mail: jbass@partners.org

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1688-4276) Copyright © 2007 by the American Academy of Pediatrics

PEDIATRICS Volume 119, Number 3, March 2007 611
Palliative care is both a philosophy of care and an organized system for delivering care. The goals of palliative care are to prevent and relieve suffering and support the best possible quality of life for patients and their families. Contrarily to common belief, palliative care is not synonymous with hospice care and should be introduced early in the disease process, possibly even at the time of diagnosis. Indeed, palliative care, with its emphasis on symptom management and quality of life, nurtures a sense of well-being and is an important coping principle in the care of children with life-threatening illnesses.

Although many strides have been made in the management of children with sickle cell disease (SCD), the life expectancy of these patients is still significantly decreased compared with the general population. Indeed, a recent study reported that half of all patients with sickle cell anemia only survive into their 40s. The chronicity of illness in SCD is punctuated by painful episodes, end-organ disease with the development of chronic sequelae, and fear of sudden death. Infection, acute chest syndrome, stroke, and multiorgan failure continue to be the main causes of death. As such, one can assume that the principles of palliative care are relevant and should be applied to patients with SCD. Current literature on pediatric palliative care, however, does not consistently list SCD as an illness that is appropriate for palliative care, and children with SCD are only occasionally included in palliative care programs once the onset of complicating sequelae has occurred.

In contrast to the palliative care philosophy in which partnership between patients and health care providers is visible, the sickle cell experience is noted for adversarial relationships with health care providers. Often times, these patients feel that the medical community is less than compassionate and is, at best, tolerant. Patients with SCD may experience emotional marginalization, amplification in depressive moods, and expressed helplessness and hopelessness. Deficiencies in school readiness have been noted in children with SCD that do not seem to be caused by the underlying medical pathology. These biopsychosocial factors can culminate in older adolescents as social anxiety and manifest as anger and low expectations; feelings of disfranchisement will result in limited productivity in the adult workforce with resultant dwindling expectations from medical and lay communities. Indeed, attitudes of the caregiver toward patients with SCD have been reported to contribute to the undertreatment of sickle cell–related pain. Medical providers are often reported to be apprehensive of the treatment requests of patients with SCD. This apprehension stems from misinterpretation of opioid requests by these patients as drug-seeking behavior. For example, a recent survey indicated that 53% of emergency department physicians and 23% of hematologists think that 20% of patients with SCD are addicts. These beliefs and perceptions about SCD-related pain and the presumed prevalence of addiction in this patient population must be addressed with objectivity if clinical care is to be changed substantively.
We strongly suggest that palliative care for patients with SCD should be introduced in early childhood. Using palliative care philosophies in the management of patients with SCD would change the public and medical purview of this population’s plight and is likely to have a tremendous impact on the health care system as well as patients and their families. Indeed, it is predicted that utilization of emergency department services alone would plummet if caregivers fostered better psychosocial adjustment and acceptance of this population’s dilemma.

An integrative medicine approach to palliative care is proposed for improving the health and well-being of patients with SCD. Integrative medicine is “the practice of medicine that reaffirms the importance of the relationship between practitioner and patient, focuses on the whole person, is informed by evidence, and makes use of all appropriate therapeutic approaches, health care professionals and disciplines to achieve optimal health and healing.” Transfer of the integrated palliative philosophy and practice to the care of pediatric patients with SCD would change the medical culture and improve this patient population’s life experience. Indeed, from the patient’s perspective, a sense of well-being would likely develop. Expressions of care, respect for patients’ problems, and the development of a therapeutic alliance with these patients would maximize patient and, ultimately, physician satisfaction.

It is well established that undertreatment of pain compromises patient autonomy and that good pain control can help preserve autonomy. Excellent access to treatment can be better assured by formal pain management education of the health care providers to impact attitudes and practices. Disease sequelae or coexisting disease place constraints on the pharmacopoeia available to treat pain in this population. For example, the triptans have a >75% success rate in the management of migraine headaches, yet this drug class is relatively contraindicated for those patients with SCD and a history of migraine headaches. Use of second-line multidrug therapy of nonsteroidal antiinflammatory drugs and opioids typically results in <50% satisfaction for migraine management. Tolerance to opioids and analgesic rebound from chronic nonsteroidal antiinflammatory drugs are potential complications for the sickle cell migraineur. Thus, we also submit that in addition to traditional allopathic therapies, complementary and alternative medicine (CAM) therapies have to be used as adjuncts for pain and symptom management. These therapies include massage, healing-touch therapies, acupuncture, and hypnosis. Hypnosis has shown promise as an effective adjunctive therapy in reducing vaso-occlusive crisis frequency and intensity. If the CAM model of palliative care was applied to the sickle cell population throughout the continuum of care, then preemptive care may result in a reduction of the severity of the disease. As of October 2006, the National Institutes of Health Clinical Center was recruiting patients 18 years and older with SCD and a history of pain for trials in hypnosis as a randomized, controlled, single-crossover, single-blinded pilot study. This investigation will examine whether hypnosis can reduce the frequency and intensity of pain in patients with SCD. Acupuncture has been used successfully in pain management when conventional therapies have failed in the treatment of vaso-occlusive crisis in hemoglobin SS/SC patients. Results of a University of Florida study suggested that body work (ie, both massage therapy and progressive muscle relaxation with guided imagery) show promise as adjunctive interventions for reducing acute and chronic SCD-related pain. The illness management of SCD is most befitting for the palliative care model with CAM adjunctive therapies.

In conclusion, the American Academy of Pediatrics recommends the development of policies and standards that promote the welfare of infants and children living with life-threatening or terminal conditions and their families, with the goal of providing equitable and effective support via an integrated model of palliative care in which the components of palliative care are offered at diagnosis and continued throughout the course of illness regardless of whether the outcome ends in cure or death. Failure to involve patients with SCD in decision-making about their care undermines their self-reliance and self-knowledge. It also reduces capacity for self-management and weakens the very fabric of health care. We strongly urge the medical establishment and our colleagues to adopt and implement the palliative care approach to children who suffer from SCD.

REFERENCES
UNION DISRUPTS PLAN TO SEND AILING WORKERS TO INDIA FOR CHEAPER MEDICAL CARE

“Carl Garrett, a 60-year-old North Carolina resident, was packing his bags to fly to New Delhi and check into the plush Indraprastha Apollo Hospital to have his gall bladder removed and the painful muscles in his left shoulder repaired. Mr Garrett was to be a test case, the first company-sponsored worker in the United States to receive medical treatment in low-cost India. But instead of making the 20-hour flight, Mr Garrett was grounded by a stormy debate between his employer, which saw the benefits of using the less expensive hospitals in India, and his union, which raised questions about the quality of overseas health care and the issue of medical liability should anything go wrong. ‘I was looking forward to the adventure of being treated in India,’ Mr Garrett said the other day. ‘But my company dropped the ball.’ The union, the United Steelworkers, stepped in after it heard about Mr Garrett’s plans, saying it deplored a ‘shocking new approach’ of sending workers to low-cost countries as a way to cut health care costs. Its officials insisted that Mr Garrett be offered a health care option within the United States.”

Noted by JFL, MD
COMMENTARY

The Use of Amplitude-Integrated Electroencephalography: Beware of Its Unintended Consequences

John M. Freeman, MD

John M. Freeman Pediatric Epilepsy Center, Department of Neurology, Johns Hopkins Medical Institutions and the Johns Hopkins Berman Bioethics Institute, Baltimore, Maryland

The author has indicated he has no financial relationships relevant to this article to disclose.

IN THE preface to the 2006 issue on brain monitoring in the neonate, White and Spitzer1 stated that continuous electroencephalogram monitoring “will substantially enhance our understanding of neonatal neurologic injury and markedly improve outcomes for all hospitalized newborn infants.” However, they caution that “access to information often precedes the ability to use it wisely.” In light of the experience of the obstetrical community with electronic fetal monitoring (EFM),2 widespread use of the new amplitude-integrated electroencephalography (aEEG) to diagnose and manage neonatal seizures should be implemented cautiously.

ELECTRONIC FETAL HEART RATE MONITORING

EFM was introduced into widespread clinical practice in the early 1960s in an effort to identify a fetus undergoing stress or distress who might develop hypoxic-ischemic encephalopathy and subsequent cerebral palsy.3,4 At the time of its introduction, there were no controlled studies to show that it was preferable to auscultation. Subsequent studies and meta-analyses have shown that although the cesarean-section rate has increased by 40% since the introduction of EFM, the perinatal morbidity and mortality rates have not changed, nor has the incidence of subsequent cerebral palsy.5–7 However, EFM has resulted in a marked increase in maternal morbidity associated with the cesarean sections. Nelson7 noted that three fourths of children who developed cerebral palsy showed no abnormalities on fetal monitoring, and monitoring had a false-positive rate of 99.8%. There has also been a dramatic increase in the incidence of malpractice suits against obstetricians, often based on a retrospective interpretation of the monitoring strips. Knowing of the adverse outcome, lawyers and experts review the child’s records and claim: “If only the delivery had been implemented earlier . . .”

Despite many recent articles analyzing these effects,2 despite knowledge of fetal oxygen saturation,8 and despite the widespread conclusion that “[o]perative intervention based on fetal monitoring may do more harm than good,”9 EFM remains the standard of care.

Implementation of aEEG monitoring in nurseries to detect and manage neonatal seizures, without clear evidence of either benefit or effective therapy of the encephalopathy detected, may lead to similar adverse medical and legal outcomes.

NEONATAL SEIZURES AND CONTINUOUS EEG MONITORING

“Seizures in the newborn are the most common and important sign of acute neonatal encephalopathy; are a major risk for death or subsequent neurologic disability; and of themselves may contribute to adverse neurodevelopmental outcome.”10 Because clinical seizures may be missed or misinterpreted by the staff, and the interpretation of neonatal EEGs is labor intensive and re-

Abbreviations: EFM, electronic fetal monitoring; EEG, electroencephalography; aEEG, amplitude-integrated electroencephalography; ENS, electroencephalogaphically detected neonatal seizure

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.2006-3650
doi:10.1542/peds.2006-3650

Accepted for publication Dec 21, 2006
Address correspondence to John M. Freeman, MD, 1026 Rolandvue Rd, Towson, MD 21204. E-mail: jfreeman@jhmi.edu

PEDIATRICS (ISSN Numbers: Print, 0031–4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
quires considerable training, the equipment for the automated detection of spikes (aEEG) is increasingly used. The arguments for its use are similar to the arguments originally used for the implementation of EFM.

Although studies in infants have not disentangled the consequences of the underlying etiology of the seizures from the additional consequences of the seizures themselves, studies in animals and neonates have suggested that clinical seizures may (but only may) have adverse effects. The consequences of subclinical seizures and electroencephalographically detected neonatal seizures (ENSs) have not been studied, and their importance has not been established.

There is consensus that if treatment of clinical seizures is to be effective, early detection is imperative—and already possible using standard crib-side observations. aEEG only slightly enhances the identification of infants at risk for subclinical seizures. However, the occurrence of 2 recent international conferences on brain monitoring in the neonate and the presence of 3 companies that were selling aEEG machines at the recent Child Neurology Society meeting (October 2006) testify to the growing interest in this technology.

Recognizing the need to establish a safe and effective treatment for neonatal seizures, Clancy, writing for the Neurology Group on Neonatal Seizures, proposed a multicenter study of the efficacy of phenobarbital in the treatment of subclinical neonatal seizures and focused on newborns “at high risk for developing early subclinical electroencephalographically detected neonatal seizures (ENS).” It is acknowledged, however, that there are no data on intraobserver reliability in reading these studies; little evidence that the detected EEG abnormalities cause problems; and minimal evidence that treatment with phenobarbital or any other anticonvulsant is of more benefit than harm.

As a study of the efficacy of phenobarbital on ENS, this multicenter project may be justifiable. However, the more widespread, routine use of this new, unproven, and unstandardized technology to detect and potentially treat clinical or subclinical neonatal seizures is fraught with hazard.

Neonatal seizures are virtually always the result of an existing or preexisting encephalopathy. The encephalopathy may be the result of clinical or subclinical infections in the mother or the infant or acute insults (such as trauma, bleeds, or strokes) and may, occasionally, be caused by fetal hypoxia or ischaemia. Whatever the etiology, there is no clear evidence that subclinical seizures requiring aEEG detection cause damage above that caused by the underlying encephalopathic agent(s). There is also no clear evidence that anticonvulsant medications can effectively control the seizures and no evidence that such treatment decreases the subsequent sequelae such as cerebral palsy or retardation. However, there is evidence that the treatments themselves could have adverse effects on the central nervous system of the newborn.

aEEG should not enjoy widespread use for the detection and treatment of neonatal seizures until it is:

1. proven useful in reliably detecting abnormalities such as subtle (subclinical) seizures;
2. shown that the test can be interpreted reliably and reproducibly by those who will use it;
3. demonstrated that the seizures identified solely by aEEG lead to neurologic dysfunction; and
4. shown that treatment prevents or ameliorates the dysfunction.

CONCLUSIONS

aEEGs are increasingly used in nurseries to detect both seizures and subclinical seizures. It has been claimed that “every child admitted to a NICU with NE [neonatal encephalopathy] or seizures should be placed on continuous monitoring” despite the lack of evidence for the deleterious effects of subclinical seizures detected by ENS. Greisen pointed out that although subtle seizures and spikes are not normal, the benefits of intervention are unclear, and the additional benefits of monitoring are also unclear. Indeed, even the importance of interictal spikes remains in doubt.

In light of the experience with the introduction of fetal monitoring and its medical-legal consequences, aEEG should be introduced into routine use with caution until evidence of its usefulness in seizure detection and the usefulness of seizure treatment is forthcoming.

REFERENCES

11. Holmes GL, Gairsa JL, Chevassus-Au-Louis N, Ben-Ari Y. Con-

ADVOCACY OVERLOAD? PROLIFERATION OF PATIENT GROUPS PROMPTS ACTIVISTS TO PUSH FOR MORE UNIFIED EFFORTS

“At the National Brain Tumor Foundation, executive director Rob Tufel has a standard reply for well-meaning families and patients who want to set up a patient-advocacy group: ‘Please don’t start another organization.’ There are 141 patient-advocacy groups that cover brain tumors, according to Mr Tufel, while 43 000 people in the United States are diagnosed every year with primary brain tumors, benign or malignant. That’s roughly one group for every 305 new patients. ‘It just doesn’t make sense from the point of view of funding, or from the point of view of patients and families,’ who must sort through the numerous organizations and Websites for information, Mr Tufel says. ‘Competition is good because it keeps us on our toes, but at some point . . . it becomes ineffective.’ Competition is an issue that many patient-advocacy groups wrestle with as they struggle to raise funds for research, attract greater interest in their diseases, and speed up the search for a cure. Increasingly, advocates are asking: When it comes to a disease like cancer, is it possible to set a common agenda and speak with a unified voice? Now some groups have begun to debate whether the large number of organizations may be hindering as much as helping in their efforts.”

Noted by JFL, MD
Endorsed Policy Statement

Cardiovascular Risk Reduction in High-Risk Pediatric Populations


All statements of endorsement from the American Academy of Pediatrics automatically expire 5 years after publication unless reaffirmed, revised, or retired at or before that time.

Executive Summary of an American Heart Association Scientific Statement: Cardiovascular Risk Reduction in High-Risk Pediatric Patients

While atherosclerosis has been clearly shown to begin in childhood, the process is usually subclinical, the rate of progression is slow, and the appropriate therapeutic approach is preventive. By contrast, certain pediatric disease states are associated with dramatically accelerated atherosclerosis, with clinical coronary events occurring in childhood or very early adult life. Intensive cardiovascular risk reduction is of critical importance in such children. This executive summary summarizes the work of an expert panel convened by the American Heart Association to develop recommendations for cardiovascular risk management in high-risk pediatric settings. The recommendations were peer reviewed and then endorsed by the American Academy of Pediatrics; the complete scientific statement was published in the December 12, 2006, issue of Circulation (2006;114:2710–2738).

In pediatric populations, a large and growing knowledge base documents the presence of accelerated atherosclerosis, the relationship of the atherosclerotic process to the number and intensity of defined risk factors, and the response at the clinical, pathologic, and vascular level to risk factor change. The panel reviewed all the available science regarding very early atherosclerotic disease as well as the range of approaches to risk assessment and treatment and the response to intervention. From this evidence, 8 pediatric disease settings were selected for inclusion: (1) familial hypercholesterolemia; (2) diabetes mellitus, type 1 and type 2; (3) chronic kidney disease; (4) post–heart transplantation; (5) Kawasaki disease; (6) chronic inflammatory disease; (7) congenital heart disease; and (8) childhood cancer survivors. Based on the presence of manifest atherosclerotic disease in childhood, a stratification protocol was developed, and each disease was classified (Table 1):
Tier I. Pathologic and/or clinical evidence for manifest coronary disease before 30 years of age;

Tier II. Pathophysiologic evidence for arterial dysfunction indicative of accelerated atherosclerosis before 30 years of age;

Tier III. Increased cardiovascular risk factors with epidemiologic evidence for coronary disease early in adult life but after 30 years of age.

Recommendations for cardiovascular risk management for each tier were tailored to the specific disease setting and adjusted for risk intensity. For children at the
TABLE 2  Tiers I, II, and III: Treatment Recommendations

**Growth/diet**
- Nutritionist evaluation, diet education for all: total fat <30% of calories, saturated fat <10% of calories, cholesterol <300 mg/d, avoid trans fats; adequate calories for growth.
- Calculate BMI percentile for sex/height.α
  - If initial BMI >95th percentile:
    - Step 1:
      - Age-appropriate reduced-calorie training for child and family
      - Specific diet/weight F/U every 2 to 4 weeks for 6 months; repeat BMI calculation at 6 months
      - Activity counseling (see below)
  - If F/U BMI >85th percentile for tier I, >90th percentile for tier II, or >95th percentile for tier III:
    - Step 2:
      - Weight-loss program referral plus exercise training program appropriate for cardiac status

**Blood pressure (tiers I, II, and III)**
- BP measurement/interpretation for age/sex/height
  - If SBP and/or DBP = 90th to 95th percentile or BP >120/80 mm Hg (3 separate occasions within 1 month):
    - Step 1: decreased calorie intake, increased activity for 6 months
  - If initial SBP and/or DBP >95th percentile (confirmed within 1 week) or 6-month F/U SBP and/or DBP >95th percentile:
    - Step 2: initiate pharmacological therapy per Fourth Task Force recommendations

**Lipids**
- LDL-C (tiers II and III)
  - See Table 3 for recommendations for LDL-C for tier I.
  - If initial LDL-C ≥130 mg/dL (tier II) or >160 mg/dL (tier III):
    - Step 1: nutritionist training for diet with <30% of calories from fat, <7% of calories from saturated fat, cholesterol <200 mg/dL, avoidance of trans fats for 6 months
  - If repeat LDL-C >130 mg/dL in tier II or >160 mg/dL in tier III and child >10 y old:
    - Step 2: initiate statin therapy with LDL goal of 130 mg/dL

**Triglycerides**
- If initial TG = 150 to 400 mg/dL:
  - Step 1:
    - Nutritionist training for low simple carbohydrate, low-fat diet
    - If elevated TGs are associated with excess weight, nutritionist referral for weight loss management: energy balance training plus activity recommendations (see below)
  - If TG >700 to 1000 mg/dL, initial or F/U:
    - Step 2:
      - Consider fibrate or niacin if >10 y old.b
      - Weight loss recommended when TG elevation is associated with overweight/obesity.

**Glucose (tiers I, II, and III, except for patients with diabetes mellitus)**
- If fasting glucose = 100 to 126 mg/dL:
  - Step 1: reduced-calorie diet, increased activity aimed at 5% to 10% decrease in weight over 6 months
- If repeat fasting glucose = 100 to 126 mg/dL:
  - Step 2: insulin-sensitizing medication per endocrinologist
- Casual glucose >200 mg/dL or fasting glucose >126 mg/dL = diabetes mellitus → endocrine referral for evaluation and management
- Maintain HbA1c <7%

**Smoking (tiers I, II, and III)**
- Step 1: parental smoking history at every visit; child smoking history beginning at age 10. Active antismoking counseling for all; smoke-free home strongly recommended at each encounter.
- Step 2: smoking cessation referral for any history of cigarette smoking.

**Activity (tiers I, II, and III)**
- For children in all tiers, participation in activity is at the discretion of the physician(s) directing care. For specific cardiac diagnoses such as Kawasaki disease and congenital heart disease, activity guidelines are referenced.
- Step 1: specific activity history for each child, focusing on time spent in active play and screen time (television + computer + video games). Goal is ≥1 hour of active play per day; screen time limited to ≤2 hours/day.
- Encourage activity at every encounter.
- Step 2: after 6 months, if goals not met, consider referral for exercise testing, recommendations from exercise specialist.

*Specific treatment goals for each risk factor and each tier are given in the algorithm (Fig 1). F/U indicates follow-up; BP, blood pressure; SBP, systolic blood pressure; DBP, diastolic blood pressure; LDL-C, low-density lipoprotein cholesterol; TG, triglycerides; HbA1C, hemoglobin A1C.*

α Normal BMI values for age and sex are available at www.cdc.gov/growthcharts.

b Elevation of triglycerides to ≥1000 mg/dL is associated with significant risk for acute pancreatitis. A fasting TG of 700 mg/dL is likely to rise to >1000 mg/dL postprandially. Treatment recommendation is congruent with guidelines for management of dyslipidemia in diabetic children.
highest risk (tier I), the intervention strategy regards the diagnosis as a “coronary heart disease equivalent” with recommendations for risk reduction similar to secondary prevention guidelines for adults with established coronary disease. For tier II, complete risk factor assessment is recommended with specific defined therapeutic goals. For children with diagnoses in tier III, the focus is on complete risk factor assessment with therapeutic goals as defined for children in general.

Recommendations for evaluation and treatment are summarized in a treatment algorithm (Fig 1) and in 2 supporting tables (Tables 2 and 3). For review of the evidence for early coronary disease and the response to intervention as well as supporting references, readers are referred to the complete scientific statement.

Further research is needed to explore the pathophysiology of atherosclerosis unique to each of these diagnoses and to critically evaluate therapeutic interventions. Because the time course to clinical disease is short, disease settings like these offer a unique opportunity in pediatric cardiovascular research to perform prospective randomized trials of the efficacy and safety of interventions.

The recommendations presented here are directed toward the primary care providers and pediatric subspecialists who care for these patients in childhood as well as to the internists, family practitioners, and adult subspecialists who will assume their care when they reach adult life. As new information develops, the guidelines will need to be modified to improve guidance on cardiovascular risk reduction in such high-risk pediatric settings. Finally, decisions on the management of individual patients must be tailored to their unique circumstances.

### TABLE 3  Tier I Conditions: Specific Treatment Recommendations

- Rigorous age-appropriate education in diet, activity, and smoking cessation for all
- Specific therapy as needed to achieve BP, LDL-C, glucose, and HbA1c goals as indicated for each tier, as outlined in algorithm; timing individualized for each patient and diagnosis. Step 1 and step 2 therapy for all outlined in Table 2.

#### Homozygous FH
- LDL management: scheduled apheresis every 1 to 2 weeks beginning at diagnosis to maximally lower LDL-C, plus statin and cholesterol absorption inhibitor  
  - Rx per cardiologist/lipid specialist. (Specific therapeutic goals for LDL-C are not meaningful with this diagnosis.)
  - Assess BMI, BP, and FG: step 1 management for 6 months
  - If tier I goals not achieved, proceed to step 2.

#### Diabetes mellitus, type 1
- Intensive glucose management per endocrinologist, with frequent glucose monitoring/insulin titration to maintain PG < 200 mg/dL, HbA1c < 7%
  - Assess BMI, fasting lipids: step 1 management of weight, lipids for 6 months
  - If goals not achieved, proceed to step 2; statin Rx if >10 y old to achieve tier I treatment goals
  - Initial BP > 90th percentile: step 1 management plus no added salt, increased activity for 6 months
  - BP consistently >95th percentile for age/sex/height: initiate ACE inhibitor therapy with BP goal <90th percentile or <130/80 mm Hg, whichever is lower.

#### CKD/end-stage renal disease
- Optimization of renal failure management with dialysis/transplantation per nephrology
  - Assess BMI, BP, lipids, FG: step 1 management for 6 months
  - If goals not achieved, proceed to step 2; statin Rx if >10 y old to achieve tier I treatment goals

#### After heart transplantation
- Optimization of antirejection therapy, treatment for CMV, routine evaluation by angiography/perfusion imaging per transplant physician
  - Assess BMI, BP, lipids, FG: initiate step 2 therapy, including statins, immediately in all patients >1 y old to achieve tier I treatment goals

#### Kawasaki disease with coronary aneurysms
- Antithrombotic therapy, activity restriction, ongoing myocardial perfusion evaluation per cardiologist
  - Assess BMI, BP, lipids, FG: step 1 management for 6 months
  - If goals not achieved, proceed to step 2; statin Rx if >10 y old to achieve tier I treatment goals

BP indicates blood pressure; LDL-C, low-density lipoprotein cholesterol; HbA1c, hemoglobin A1c; FH, familial hypercholesterolemia; Rx, prescription/treatment; FG, fasting glucose; PG, plasma glucose; ACE, angiotensin-converting enzyme; CKD, chronic kidney disease; CMV, cytomegalovirus.
High-Deductible Health Plans and the New Risks of Consumer-Driven Health Insurance Products

Committee on Child Health Financing

ABSTRACT
Consumer-driven health care is the most noteworthy development in health insurance since the widespread adoption of health maintenance organizations and preferred provider organizations in the 1980s. The most common consumer-driven health plan is the high-deductible health plan, which is essentially a catastrophic health insurance plan, often linked with tax-advantaged spending accounts, with very high deductibles, fewer benefits, and higher cost-sharing than conventional health maintenance organization or preferred provider organization plans. The financial risks are significant under high-deductible health plans, especially for low- to moderate-income families and for families whose children have special health care needs. Of concern for pediatricians are the potential quality risks that are predictable in high-deductible health plans, in which families are likely to delay or avoid seeking care, especially preventive care (if it is not exempted from the deductible), when they are faced with paying for care before the deductible is met. This policy statement provides background information on the most common consumer-driven health plan model, discusses the implications for pediatricians and families, and offers recommendations pertaining to health plan product design, education, practice administration, and research.

INTRODUCTION
Consumer-driven health care is the most noteworthy development in health insurance since the widespread adoption of health maintenance organizations (HMOs) and preferred provider organizations (PPOs) in the 1980s. Faced with unsustainable premium increases and heightened competition, employers are experimenting with new products, referred to as consumer-driven health plans (CDHPs). The potential benefit of a CDHP is to increase the control consumers have over their health care spending and to empower them to use published information to guide their care options. The most commonly sold CDHP is a high-deductible health plan (HDHP), which essentially is a catastrophic health insurance plan, often linked with tax-advantaged spending accounts, with very high deductibles, fewer benefits, and higher cost-sharing than conventional HMO and PPO plans. HDHPs offer a new strategy for sharing risk and responsibility for health care costs among employers and employees. HDHPs also represent a major shift from defined benefits to defined contributions. At this time, there is insufficient information to ascertain the specific effects of HDHPs on children’s access to care and the operation of the medical home; however, there is concern that...
children from low- to moderate-income families and children with special health care needs may be at risk if covered under HDHPs.

This policy statement provides background information on the most common CDHP model—the HDHP paired with a tax-advantaged spending account—and the latest research on these new insurance products. The statement also discusses the implications for pediatricians and families and offers recommendations pertaining to product design, education, practice administration, and research.

**BACKGROUND**

HDHPs were established by the Medicare Prescription Drug Improvement and Modernization Act of 2003. Currently, qualified HDHPs are health insurance plans with at least a $2000 deductible for family coverage and a total annual out-of-pocket maximum, including deductible, copays, and other cost-sharing, that cannot exceed $10,000 per family. Money can be carried over from year to year, and the spending account—either a health reimbursement account (HRA) or a health savings account (HSA)—is used to pay for a portion of health care expenses until the plan’s high deductible is met; then, the HDHP functions like a PPO plan.

The 2 common spending accounts differ in terms of ownership, requirements to be tied to an HDHP, discretion to carry over unused amounts into subsequent years, and portability (see Table 1). Briefly, the HRA is owned and solely funded by the employer. It is typically offered with an HDHP but also can be offered with an HMO or PPO. The employer has discretion about the amount of funds to be carried over, and HRAs are not portable. The HSA is owned by the employee, although the employer can contribute. It can only be used with an HDHP that has a deductible up to $5150 per family. Money can be carried over from year to year, and the HSA is portable.

Research shows that very large and very small employers as well as individuals in the nongroup market are most interested in offering and purchasing HDHPs. Early research suggests that healthier and wealthier individuals are more likely to purchase HDHPs than their counterparts. Individuals and families in higher tax brackets, especially those who are healthy, can benefit from this method to save for medical expenses, and possibly retirement, with pretax dollars.

According to the 2005 National Employers Health Benefits Survey sponsored by the Kaiser Family Foundation and the Health Research and Education Trust, 20% of employers are offering an HDHP, up from only 5% in 2003. A minority of firms that offer HDHPs (1 in 5) offer either an HRA contribution (10%) or an HSA-qualified plan (12%). In firms that offer HRAs, approximately 25% of employees participate (1.6 million employees or 2% of all covered workers). On average, employer contributions to HRAs amount to $1556. In firms that offer HSAs, approximately 15% of employees participate ($10,000 employees or 1.2% of all covered workers). Average annual employer contributions to HSAs amount to $1185, with one third of employers making no contributions. It is unclear what the average amount that employees contribute is.

Table 2 illustrates the cost differences between HDHPs and PPOs for an average family. Although the proportion of employers currently offering HDHPs with spending accounts is small, the expected growth is predicted to be sizeable. According to the 2005 National Employer Health Benefits Survey, 2% to 4% of firms reported that they were very likely to offer HDHPs next year, and 22% to 25% reported that they were somewhat likely to offer them.

**TABLE 1** Comparison of HRA and HSA

<table>
<thead>
<tr>
<th>Plan</th>
<th>Tax Savings</th>
<th>Funded by</th>
<th>Annual Rollover of Unused Funds</th>
<th>Portable</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRA</td>
<td>Yes</td>
<td>Employer</td>
<td>At the employer’s discretion</td>
<td>At the employer’s discretion</td>
</tr>
<tr>
<td>HSA</td>
<td>Yes (funds may be invested and earn interest tax free)</td>
<td>Can be both employer and employee</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**PREVENTIVE CARE AND HDHPs**

Generally, an HDHP cannot provide any benefits before the deductible is satisfied, but there is an exception for preventive care. Referred to as the “safe harbor for preventive benefits,” HDHPs with HSAs are permitted, but not required, to offer preventive care without meeting the deductible. According to the Internal Revenue Code, preventive care is defined as routine well-child care and immunizations; periodic health evaluations, including tests and diagnostic procedures ordered in conjunction with routine examinations, such as annual physicals; mental health and substance abuse screening; vision and hearing screening; screenings for various pediatric conditions (ie, developmental delay, congenital hypothyroidism, lead concentration, phenylketonuria, and scoliosis); metabolic, nutritional, and endocrine screening; infectious disease screening; and maintenance drugs used by chronically ill patients. Despite this important provision, the 2005 National Employer Health Benefits Survey found that only 30% of employers who offer an HDHP with an HSA covered preventive care before the deductible was met, thus eroding the relationship between the medical home and the family.
IMPLICATIONS FOR PEDIATRICIANS AND FAMILIES
HDHPs carry potentially significant coverage, financial, quality, and practice risks for pediatricians as well as families. Among the coverage restrictions, HDHPs typically offer less generous coverage for certain services (e.g., drugs, mental health) compared with PPOs or HMOs.10 Physician and hospital coverage is likely to be the same, although not necessarily in terms of cost-sharing.8 It is often difficult to assess the coverage risks associated with preventive care, because the service may or may not be exempt from the deductible; also, information on periodicity and content may not be extensively described in consumer materials. Another more significant coverage risk of HDHPs is the potential for “destabilization” of employer-sponsored health insurance if more employers and families purchase HDHPs instead of HMOs and PPOs, which typically offer more comprehensive benefits.

The financial risks are significant under HDHPs, especially for low- to moderate-income families and for families whose children have special health care needs.7 Because visits by children with special health care needs to specialists are not considered preventive care, parents will need to tap into their HSA or HRAs to pay for these visits as well as any laboratory tests, imaging, therapies, and other essential health care services. Once the HSA or HRA is depleted, parents will need to pay out-of-pocket until they have reached their deductible. Thus, children with special health care needs may not receive all their preventive care, because the service may or may not be exempt from the deductible. In addition, families may be asked more about the costs of their services as well as the content and value of specific services.13 In addition, families in these plans will likely request more telephone and e-mail assistance to avoid making in-person visits.

Of concern to pediatricians are the potential quality risks that are predictable in HDHPs in which families are likely to delay or avoid seeking care when they are faced with paying for care before the deductible is met.14 Lower rates of preventive care and immunizations, less compliance with recommended treatment, less continuity of care, and lower use of acute and chronic care services are very real concerns.15 HDHPs have the potential to adversely decrease access to medical homes and result in more episodic, high-priced care. Faced with difficult choices, families may seek to “load up” on a scheduled visit to save money or delay care until after the deductible is met. In the end, families will have to make many more decisions about the cost-versus-quality trade-offs, relying on Internet-based information, online patient support tools, and nurse help lines.1

Although decision-support tools have been identified as a special feature of HDHPs, a recent US Government Accountability Office report16 noted that tools provided by insurance carriers to assist consumers in assessing the price and quality of health care providers and services do not provide sufficient information to allow enrollees to fully assess the cost and quality trade-offs of health care–purchasing decisions. Of concern are the methods that insurers and third-party agencies use to rate the quality of care of providers. Relying on claims data, for example, represents a flawed approach to judging quality.

A variety of pediatric practice risks are starting to emerge with HDHPs. Among them are greater administrative and collection costs and bad debt for practices.17 This is attributable in part to the fact that some HDHP administrators have notified families not to pay the physician charges at the time of service, instead waiting for explanation of benefit statements to assess deductibles and savings account balances. Importantly, pediatricians are likely to be asked more about the costs of their services as well as the content and value of specific services.13 In addition, families in these plans will likely request more telephone and e-mail assistance to avoid making in-person visits.

RECOMMENDATIONS
The following recommendations focus on the different groups of people and organizations affected by CDHPs. These groups include the insurance companies and third-party payers designing the plans, the families purchasing the plans, and the employers providing the plans. Also included are recommendations for physicians and practices to prepare for CDHPs.

HDHP Design

- Coverage should be provided for preventive services including, but not limited to, well-child care, immunizations, and appropriate screenings.
• Preventive services should be “first-dollar” coverage (ie, covered before the deductible is met).
• Allowed reimbursement amounts for preventive services should be age adjusted to provide adequate payment for preventive health care recommended by the American Academy of Pediatrics (AAP).
• Physicians should be allowed to collect copays and payment for nonpreventive services at the time of visit. Methods to make this simpler, such as real-time debit cards for HSAs, should be developed. Vendors should implement integrated, real-time claims-adjudication processes to help clinicians obtain payment for services from the patient more quickly.
• Payment for services before the deductible has been met should be at billed charges. If a contracted fee schedule is used, it should be adjusted to reflect the increased billing and administration costs incurred by the physician.
• Consideration should be given to payment for telephone and e-mail services, because telephone and e-mail advice will be in greater demand.

Education
• Increase pediatricians’ awareness of the prevalence of HDHPs in their geographic area and their varied cost-sharing requirements and benefit designs.
• Communicate to employers the importance of covering preventive care outside the deductible and the importance of receiving preventive care in the medical home.
• Publicize to employers, patients, and the public the average costs of preventive care services, including the increased frequency of examinations and number of vaccines required during the first 2 years of life and the increased amount of time required for adolescent care.
• Consider new educational strategies to assist families when insurance decisions are made, and focus on deductible levels, preventive care coverage, cost-sharing protections, provider networks, spending accounts, and payment arrangements.

Practice Management
• Publicize the practice’s policy about collecting payment for services at the time of the visit.
• Communicate the costs and reasons for preventive, acute, follow-up, and chronic medical care.
• Establish billing policies for telephone and e-mail services.
• Prepare for greater administration/collection burdens and bad debts.
• Use AAP Hassle Factor forms (available online at the Member Center at www.aap.org/moc [under “more resources”]) to inform state and national AAP leaders of issues and problems.

Quality Improvement Measures
• HDHPs should adhere to providing quality data information to consumers on the basis of measurement standards developed by accrediting organizations.
• Quality data should be based on measures that are evidence based, relevant to patient outcomes, and statistically valid and reliable.

Research
• Encourage, support, and promote research to assess the value and benefits of preventive pediatric services and promote research to evaluate the effects that HDHPs have on children’s and adolescents’ access to care and family satisfaction with care and cost of care.
• Examine the effect of HDHPs on the use of medical services, including preventive, acute, and chronic care.

CONCLUSIONS
CDHPs offer the opportunity of more consumer involvement in the decision to purchase specific health care items. However, of notable concern are the effects of this process on children receiving necessary and highly cost-effective preventive care and on lower- or middle-income parents, who will have to pay for a substantial amount of their children’s health care out-of-pocket.

COMMITTEE ON CHILD HEALTH FINANCING, 2005–2006
Thomas K. McInerny, MD, Chairperson
Charles J. Barone, II, MD
Anthony D. Johnson, MD
Richard Lander, MD
Richard Y. Mitsunaga, MD
Mark S. Reuben, MD
Corrine A. Walentik, MD
Steven E. Wegner, MD, JD
Mark J. Werner, MD, CPE

CONSULTANT
Margaret McManus, MHS

STAFF
Teri Salus, MPA

*Lead authors

REFERENCES
2. Kuraitis V, Riedel J. CDHPs + DM = population health? Pre-
sented at: Disease Management Association of America Disease
Management Leadership Forum; October 17, 2005; San Diego,
CA
3. Gabel JR, Lo Sasso AT, Rice T. Consumer-driven health plans:
are they more than talk now? Health Aff (Millwood). 2002;
(suppl Web exclusives):W395–W407
4. Goff V. Consumer cost sharing in private health insurance: on
the threshold of change. Issue Brief Natl Health Policy Forum.
2004;(798):1–19. Available at: www.nhpf.org/pdfs/ib798-
5. Gabel J, Rice T. Insurance Markets: Understanding Consumer-
Directed Health Care in California. Oakland, CA: California
Available at: www.chcf.org/documents/insurance/Consumer
6. American Academy of Actuaries. The Impact of Consumer-Driven
Health Plans on Health Care Costs: A Closer Look at Plans With
Health Reimbursement Accounts. Public Policy Monograph. Wash-
at: www.actuary.org/pdf/health/cdhp/jan04.pdf. Accessed May 1,
2006
7. McManus P. Consumer-driven health care. Presented at:
American Academy of Pediatrics National Conference &
Exhibition; October 9, 2005; Washington, DC
8. Kaiser Family Foundation, Health Research and Education
Trust. Employer health benefits 2005 annual survey. Available at:
www.kff.org/insurance/chcm091405nr.cfm. Accessed May 1,
2006
223(c)(2). Available at: www.ustreas.gov/press/releases/reports/
notice200423.pdf. Accessed May 1, 2006
Health Financ Manage. 2003;57(9):46–48
11. Christianson JB, Parente ST, Feldman R. Consumer experi-
ences in a consumer-driven health plan. Health Serv Res.
2004;39(4):1123–1140
care: solution or stopgap measure? Presented at: the National
Business Coalition on Health National Health Leadership
Council Meeting; August 13–15, 2003; San Diego, CA. Avail-
May 1, 2006
13. Gauthier AK, Clancy CM. Consumer-driven health care: be-
yond rhetoric with research and experience. Health Serv Res.
14. Parente ST, Feldman R, Christianson JB. Evaluation of the
effect of a consumer-driven health plan on medical care ex-
penditures and utilization. Health Serv Res. 2004;39(4):
1189–1210
15. Mulligan K. Consumer-driven health plans not total answer to
pn.psychiatryonline.org/cgi/content/full/39/18/8. Accessed
May 1, 2006
Plans: Small but Growing Enrollment Fueled by Rising Cost of Health
Office; 2006
17. Hummel J, Singh SP. Preparing for consumerism: the top 10—
health plans and providers alike should prepare for consumer-
driven health plans with this list of 10 strategic imperatives.
Health Manag Technol. 2004;25(10):60, 62, 64
POLICY STATEMENT

Testing for Drugs of Abuse in Children and Adolescents: Addendum—Testing in Schools and at Home

Committee on Substance Abuse and Council on School Health

ABSTRACT

The American Academy of Pediatrics continues to believe that adolescents should not be drug tested without their knowledge and consent. Recent US Supreme Court decisions and market forces have resulted in recommendations for drug testing of adolescents at school and products for parents to use to test adolescents at home. The American Academy of Pediatrics has strong reservations about testing adolescents at school or at home and believes that more research is needed on both safety and efficacy before school-based testing programs are implemented. The American Academy of Pediatrics also believes that more adolescent-specific substance abuse treatment resources are needed to ensure that testing leads to early rehabilitation rather than to punitive measures only.

BACKGROUND

In 1996, the American Academy of Pediatrics (AAP) published (and reaffirmed in 2006) a policy statement titled “Testing for Drugs of Abuse in Children and Adolescents,” which opposed involuntary testing of adolescents for drugs of abuse. The policy statement also stated that laboratory testing for drugs under any circumstances is improper unless the patient and clinician can be assured that the test procedure is valid and reliable and patient confidentiality is ensured. This policy statement was published shortly after a 1995 US Supreme Court ruling (Vernonia v Acton [515 US 646]) held that random drug testing of high school athletes is constitutional. Since that time, national interest in school-based drug testing has increased. In June 2002, the US Supreme Court, in a 5-to-4 decision, ruled that public schools have the authority to perform random drug tests on all middle and high school students participating in extracurricular activities (Board of Education v Earls [536 US 822, 122 S Ct 2559, 153 L Ed 2 days 735 [2002]]). Writing for the majority, Justice Clarence Thomas wrote, “Testing students who participate in extracurricular activities is a reasonably effective means of addressing the School District’s legitimate concerns in preventing, deterring and detecting drug use.” Shortly after this Supreme Court ruling, the President’s Office of National Drug Control Policy published a guidebook designed to encourage schools to incorporate drug-testing policies for all students.

Interest in drug testing of adolescents reaches beyond public schools. During recent years, a substantial number of companies have begun to market home testing products for adolescents.

Key Words

adolescence, substance-related disorders, substance abuse detection

Abbreviation

AAP—American Academy of Pediatrics

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
drug-testing products directly to parents.\textsuperscript{3} Products that identify alcohol and drugs in urine, saliva, and hair are now available at retail outlets and via the Internet. Pediatricians may be asked about home drug testing by parents of their adolescent patients. Pediatricians involved in school health may be asked to assist in implementing school-based drug-testing programs. For these reasons, the Committee on Substance Abuse has conducted a review of the available science on drug testing of adolescents and is issuing this addendum to the 1996 policy statement. Although much has been written on the pros and cons of testing adolescents for drugs, relatively little has been published in peer-review scientific journals.

**BENEFITS AND RISKS OF DRUG TESTING IN SCHOOLS AND AT HOME**

School- and home-based drug testing poses a number of potential benefits and risks. On the positive side, both procedures would likely increase the number of adolescents who are screened for use of illicit drugs. Population-based screening also offers the potential for providing early intervention and treatment services to more adolescents. The Office of National Drug Control Policy guidebook states: “Results of a positive drug test should not be used merely to punish a student. Drug and alcohol use can lead to addiction, and punishment alone may not necessarily halt this progression. However, the road to addiction can be blocked by timely intervention and appropriate treatment.”\textsuperscript{2} Proponents of drug testing also claim that the existence of a school- or home-based drug-testing program will help adolescents refuse drugs and provide legitimate reasons to resist peer pressure to use drugs, although these claims are not yet proven. On the negative side, drug testing poses substantial risks—in particular, the risk of harming the parent-child and school-child relationships by creating an environment of resentment, distrust, and suspicion.\textsuperscript{4} In addition to the effects on the individual adolescent, the safety and efficacy of random drug testing requires additional scientific evaluation. Broad implementation of random drug testing as a component of a comprehensive drug-use prevention program should await the results of these studies.

Currently, there is little evidence of the effectiveness of school-based drug testing in the scientific literature. Goldberg et al\textsuperscript{5} compared 2 schools, one of which implemented a mandatory drug-testing program for student athletes and the other of which did not. They found at follow-up that the use of illicit drugs, but not alcohol, was significantly lower among athletes who were drug tested. However, they also found that athletes who were drug tested experienced an increase in known risk factors for drug use, including an increase in normative views of use, belief in lower risk of use, and poorer attitudes toward the school.

A larger observational study by Yamaguchi et al,\textsuperscript{6} which analyzed data from the national Monitoring the Future study, found no association between school-based drug testing and students’ reports of drug use. Among the nationally representative group of more than 300 schools, drug testing was most commonly conducted “for cause” (ie, suspicion; 14% of schools) and was far less commonly required for student athletes (4.9% of schools) or students participating in other extracurricular activities (2.3% of schools). Regardless of the reason it was performed, drug testing was not significantly associated with reduction in the use of marijuana or any other illicit drug among students in any grade studied (ie, 8th, 10th, or 12th grade). However, 1 observational study is not sufficient to establish causation or lack of causation. In addition, no detail was provided regarding the extent of drug testing in the study schools, and at some schools, it may have been minimal. Further scientific investigation is warranted.

Laboratory testing for drugs of abuse is a technically complex procedure. To ensure the validity of the specimen, urination must be directly observed, which is a potentially embarrassing procedure for all involved, or the collector must use a fairly complex and expensive federally approved protocol, which involves documentation of a continuous chain of custody in handling and includes temperature testing and controls for adulteration and dilution.\textsuperscript{7} Few schools will have sufficient staff with proper training to implement these costly procedures, and a recent survey of pediatricians, adolescent medicine specialists, and family physicians found that few physicians will be able to help, because less than 25% are familiar with proper procedures for collection, validation, and interpretation of urine drug tests.\textsuperscript{8} Similarly, most parents cannot implement the federal collection protocol and, for ethical and developmental reasons, should not directly observe their teenaged children urinating. Although drug testing of hair and saliva is available, validity has not been firmly established. Questions remain regarding how passive exposure to drugs as well as differences among races and sexes can affect hair testing.\textsuperscript{9-12} In addition, hair testing is more likely to be useful in detecting historical drug use rather than current use.\textsuperscript{9,13} Oral fluid testing (ie, saliva or oral swab), by contrast, gives a more accurate picture of current use.\textsuperscript{14} However, accuracy of oral fluid testing varies across drugs of abuse. Oral fluid testing performs well in detecting the use of opiates and methamphetamine, but it performs poorly in detecting the use of benzodiazepines and cannabinoids.\textsuperscript{15-17} Interpretation of drug tests can also be complex. School staff members and/or parents need to be able to assess possible false-positive results, especially when screening test results are positive for amphetamines or opioids. Over-the-counter cold medications containing pseudoephedrine can cause false-positive screening re-
sults for amphetamine, although follow-up testing with gas chromatography and mass spectrometry is highly specific and can reliably confirm the presence of amphetamine. Ingestion of foods that contain poppy seeds makes interpretation of drug testing more difficult, because it can cause screening and gas chromatography and mass spectrometry results to be falsely positive for morphine and/or codeine.

It is fairly easy to defeat drug tests, and most drug-involved youth are all too familiar with ways to do so. Even properly collected specimens must have checks for validity (eg, urine specific gravity and creatinine), because the easiest way to defeat a drug testing is by simple dilution. Even when properly collected and validated, urine drug tests yield very limited information. With the exception of marijuana, the window of detection for most drugs of abuse is 72 hours or less. Therefore, negative test results indicate only that the adolescent did not use a specific drug during the past several days. Even adolescents with serious drug problems may have negative test results on most occasions. Standard drug-testing panels also do not detect many of the drugs most frequently abused by adolescents, such as alcohol, ecstasy (3,4-methylenedioxymethylamphetamine [MDMA]), and inhalants, and information on the limitations of screening tests and ways to defeat them is widely available to adolescents via the Internet. Widespread implementation of drug testing may, therefore, inadvertently encourage more students to abuse alcohol, which is associated with more adolescent deaths than any illicit drug but is not included in many standard testing panels. Mandatory drug testing may also motivate some drug-involved adolescents to change from using drugs with relatively less associated morbidity and mortality, such as marijuana, to those that pose greater danger (eg, inhalants) but are not detected by screening tests. No studies have yet been conducted on this important issue. Safety of randomly testing adolescents for the use of drugs should be scientifically established before it is widely implemented.

Drug testing may also be perceived by adolescents as an unwarranted invasion of privacy. A policy statement is being developed by the Council on School Health on the role of schools in combating substance abuse. It will discuss the potential risks of school-based drug testing and alternative approaches to school-based prevention of drug abuse. Few physicians support school-based testing of adolescents for drugs; a national survey of physicians (pediatrics, family medicine, and adolescent medicine) found that 83% disagreed with drug testing in public schools.

A key issue at the heart of the drug-testing dilemma is the lack of developmentally appropriate adolescent substance abuse and mental health treatment. Adequate resources for assessment and treatment must be available to students who have positive test results. However, many communities lack substance abuse treatment services dedicated to adolescents, and adult substance abuse treatment programs may be inappropriate and ineffective for adolescents. Federal support for school-based drug testing should include an allocation of resources that will facilitate greater access to adolescent substance abuse treatment.

ADDITIONAL CONCLUSIONS AND RECOMMENDATIONS
1. The AAP supports rigorous scientific study of both the safety and efficacy of school- and home-based drug testing of adolescents.
2. The AAP recommends that school- and home-based drug testing not be implemented before its safety and efficacy are established and adequate substance abuse assessment and treatment services are available.
3. The AAP encourages parents who are concerned that their child may be using drugs or alcohol to consult their child’s primary care physician or other health professional rather than rely on school-based drug screening or use home drug-testing products.
4. The AAP recommends that health care professionals who obtain drug tests or assist others in interpreting the results of drug tests be knowledgeable about the relevant technical aspects and limitations of the procedures.

COMMITTEE ON SUBSTANCE ABUSE, 2005–2006
Alain Joffe, MD, MPH, Chairperson
Mary Lou Behnke, MD
*John R. Knight, MD
Patricia Kokotailo, MD
Tammy H. Sims, MD, MS
Janet F. Williams, MD

PAST COMMITTEE MEMBER
John W. Kulig, MD, MPH

LIAISON
Deborah Simkin, MD
American Association of Child and Adolescent Psychiatry

CONSULTANTS
Linn Goldberg, MD
Sharon Levy, MD, MPH

STAFF
Karen Smith

COUNCIL ON SCHOOL HEALTH, 2005–2006
Robert D. Murray, MD, Chairperson
Barbara L. Frankowski, MD, MPH
Rani S. Gereige, MD, MPH
*Cynthia J. Mears, DO
Michele M. Roland, MD
Thomas L. Young, MD
Linda M. Grant, MD, MPH
Daniel Hyman, MD
Harold Magalnick, MD
George J. Monteverdi, MD
Evan G. Pattishall III, MD

LIAISONS
Nancy LaCursia, PhD
American School Health Association
Donna Mazyck, MS, RN
National Association of School Nurses
Mary Vernon-Smiley, MD, MPH
Centers for Disease Control and Prevention
Robin Wallace, MD
Independent School Health Association

STAFF
Madra Guinn-Jones, MPH

*Lead authors

REFERENCES
CLINICAL REPORT

Special Requirements of Electronic Health Record Systems in Pediatrics

S. Andrew Spooner, MD, MS, and the Council on Clinical Information Technology

ABSTRACT

Some functions of an electronic health record system are much more important in providing pediatric care than in adult care. Pediatricians commonly complain about the absence of these “pediatric functions” when they are not available in electronic health record systems. To stimulate electronic health record system vendors to recognize and incorporate pediatric functionality into pediatric electronic health record systems, this clinical report reviews the major functions of importance to child health care providers. Also reviewed are important but less critical functions, any of which might be of major importance in a particular clinical context. The major areas described here are immunization management, growth tracking, medication dosing, data norms, and privacy in special pediatric populations. The American Academy of Pediatrics believes that if the functions described in this document are supported in all electronic health record systems, these systems will be more useful for patients of all ages.

INTRODUCTION

Child health care providers often find that clinical information systems have limited usefulness in pediatrics,1,2 because they seem to be designed for adult care. For the purposes of this report, we use the definition of the electronic health record (EHR) system proposed by the Institute of Medicine:

“An EHR system includes (1) longitudinal collection of electronic health information for and about persons, where health information is defined as information pertaining to the health of an individual or health care provided to an individual; (2) immediate electronic access to person- and population-level information by authorized, and only authorized, users; (3) provision of knowledge and decision-support that enhance the quality, safety, and efficiency of patient care; and (4) support of efficient processes for health care delivery. Critical building blocks of an EHR system are the electronic health records (EHR) maintained by providers...and by individuals (also called personal health records).”

The definition proposed by the Institute of Medicine is functional in nature. It assumes that an EHR system must provide these features to be of value. Even for child health care providers, this definition is valid, and this set of features is likely to provide value to most practitioners. However, as has been noted previously,2 when viewed from the perspective of the child health care provider, these features may fall short either in the details of how they are implemented or by omitting functions that are more routine in pediatric care than in any other primary care practice. This report provides a look at these key functional requirements through the lens of the child health care provider and augments these requirements with...
the additional functions that child health care providers use in their daily practice of medicine. This report focuses on the clinical functions of the EHR system operated by the health care provider, as opposed to the more administrative functions in the practice-management system (such as appointment management, insurance eligibility determination, and billing). However, it is assumed that the EHR system in the pediatric setting is fully connected to the practice-management system through an appropriate interface or through software integration of the 2 systems.

PEDiatric FUNCTIONS
In 2001, the American Academy of Pediatrics (AAP) published a description of the features that would be desirable in a clinical information system to be used in pediatrics. Almost none of these features were purely pediatric. For example, that statement called for medication dosing by weight and for opportunities to record information about guardianship. There are certainly instances of medication dosing by body weight in adult medicine, and many adults have guardians. Yet, these features are vastly more important in pediatrics, so it is appropriate to refer to them as "pediatric functions." Several of these functions that are of critical importance to pediatric practice are discussed in greater detail here. Others are of less general importance but have been identified as desirable by members of the Pediatrics Data Standards Special Interest Group of Health Level Seven (HL7), an international health data standards development organization in which the AAP participates (www.hl7.org).

CRITICAL PEDIATRIC EHR FUNCTIONAL AREAS
There are some functional areas that are so critical to the care of infants, children, and adolescents that their absence results in the system impeding quality pediatric care.

Immunization Management

Recording Immunization Data
The ability to record multiple immunizations efficiently is critical for pediatric health maintenance activities. State and federal regulations add a complexity to the process of recording immunization administration that is absent for medications. Systems designed to record adult immunizations and other medications naturally allow the practitioner to record data such as the manufacturer, lot number, date, site, route of administration, and expiration date. The nature of immunization practices in children adds some requirements to this list, in particular, data required by the Vaccines for Children (VFC) program and the National Childhood Vaccine Injury Act (NCVIA) of 1986 (42 USC §§300aa-1–300aa-34). The VFC program, a federal program by which eligible children are provided vaccine at no charge, requires providers to maintain a separate stock of vaccine, to assess eligibility for the program, and to submit reports to the program. All of these activities require support from the information system used to track immunization data. The NCVIA has numerous implications for immunization data recording. Among these is the requirement to deliver to the parent (or equivalent health decision-maker) a vaccine information statement (VIS) and to record when it was given and which version of the VIS was given. The NCVIA also mandates that health care providers report adverse events associated with vaccines; although this applies equally to adult providers, automation of this reporting capability would be of particular interest to child health care providers, who give the bulk of vaccines. The Centers for Disease Control and Prevention’s National Immunization Program (www.cdc.gov/nip) specifies these information-management requirements in detail. EHR systems also need to manage the record of consent for vaccine administration. Vaccine refusal by a parent or patient requires the recording of refusal reasons and recording of which refused vaccines were offered.

Linking to Immunization Information Systems
Most states and several local jurisdictions have electronic immunization-information systems or registries. The EHR should allow interoperability with these systems, including the ability to download, upload, and synchronize a child’s immunization history. Some technical standards already exist for immunization information system functions and communications with them.

Immunization Decision Support
Systems for encoding rules about which immunizations are due and when they are projected to be due in the future have been in existence for years. For an EHR system to fully support pediatric practice, it must be able to take previous immunization data and derive, at the point of care, logical conclusions about the currency of immunization and recommend the appropriate immunizations. This functionality requires an understanding of the individual antigens present in each vaccine and analysis of when, in what form, and at what age in the child’s life each antigen was—or was supposed to be—administered. There may also be local variations in this functionality based on local epidemiology. These functions might be built into the system or be derived from immunization registries or third-party programs accessed via a network. If the logic is built into the EHR system itself, there should be a way to easily update the logic to reflect changes to immunization rules and to handle new vaccines and new antigen combinations.
Growth Tracking

Graphical Representation
Child health care providers make important judgments about a child’s health by visual inspection of a plot of a child’s body measurements (usually weight, height, head circumference, BMI) over time. Plots show the progression of measured values over time against curves of predicted growth or percentile curves. Ideally, the visual plot should be visible at the top level of an individual record or require minimal effort for viewing. The EHR system should allow the representation of percentile curves from a usual source (Centers for Disease Control and Prevention [www.cdc.gov/growthcharts]) or other sources that may provide these curves for special populations.15 The system should allow magnification (“zooming”) of the plot to allow inspection of areas of the plot in which measurements have been frequently made. Users ought to be able to derive growth-velocity data from 2 selected data points. The system should distinguish height from length. Also, the system should accommodate corrections for preterm birth in the graphical display of body measurements.

Percentile Calculations
In addition to representation of body measurements, the percentile value of any particular body measurement against a defined distribution is desirable. Such percentile values should be calculated and displayed at the time of data entry. Percentile values should also be available for decision-support functions of the EHR system.

Medication Dosing

Dosing by Body Weight
The predominant method for calculating pediatric drug dosages is to compute them on the basis of body weight. When a current body weight is available, the EHR system should be able to incorporate it into the prescribing process and suggest doses on the basis of accepted references. Failing this, the EHR system should make weight visible in all displays associated with drug dosing. When a current body weight is not available, the system should react to this appropriately by requesting its input. For medications that require adjustment of dose as the child’s weight increases, the intended dosage per unit of body weight should be recordable and maintained as an aspect of the prescription. Systems should be able to determine if a body weight obtained in the past is too old to be used in decision support (eg, last month’s weight would be appropriate for an adolescent but not a neonate). Entries of height, weight, and head circumference should be checked against age-based norms so that users can be warned of possible errors. As in adult care, medication dosing by body surface area or ideal weight should also be available; however, the equations for the estimation of body surface area and ideal body weight in children are different from those in adult care.

Dose-Range Checking
With or without dosing decision support, an EHR system should be able to check drug doses posthoc by using accepted pediatric references and advise the user when no pediatric references exist.

Rounding to Safe and Convenient Doses
Many medications for infants and young children are supplied in liquid form. Because parents and other caregivers must measure a volume of liquid for each dose of medication, child health care providers must compute a volume for each dose, round it to a convenient volume, and spend time educating caregivers on the proper volume to administer. EHR systems that facilitate prescribing should support prescriptions expressed in the volume of drug to be administered and avoid expressing the prescription solely in terms of the mass of the drug.

Age-Based Dosing Decision Support
For the case in which dosing guidelines or formulary benefits vary with age or gestational age,16 the system should incorporate those data into its decision support.

Dosing for the School Day
Pediatricians must often write prescriptions in which the medication is divided in 2 labeled packages—one for home administration and one for administration during the day at school, child care, or another care setting. EHR systems should provide the capability to generate instructions to the pharmacy to dispense a medication in this way.

Patient Identification

Newborn Identification
Although many EHR systems depend on the use of a government-issued identification number (usually the Social Security number), newborn infants do not receive these numbers for a significant period of time after birth. EHR systems should allow the registration of patients without such identifiers and allow retrieval of information on the basis of any temporary identifiers that may be used.

Prenatal Identifiers
An EHR system that allows storage of prenatal data (eg, from a fetal imaging procedure) should allow the logical connection of these data to the postnatal record once the child’s record is established in the system.

Name Changes
Infants undergo name changes because of changes in family structure or the need to change the temporary
name assigned at the birth hospital. Because clinical data are connected to the old names, EHR systems need to support retrieval of data via search on previous names.

**Ambiguous Sex**
In the case of a child with ambiguous genitalia, an EHR system ought to allow the assignment of sex as unknown and to operate normally until the sex of the patient is assigned.

**Norms for Pediatric Data**

**Numeric Data**
Norms for almost all numeric data (such as laboratory results, body measurements, scores on standardized assessments, and vital signs) change as the child grows. For many of these data, norms change continuously with age, so it is insufficient to provide merely a handful of normative ranges. Developers should assume that all numeric data collected in a pediatric context have changing norms over the lifespan and should provide ways of flagging abnormal values at any age. Percentile values and $z$ scores (number of SDs from the mean) should be available for those few data for which the distributions are known, such as height, weight, head circumference, and BMI.

**Nonnumeric Data**
Whenever an EHR system distinguishes normal from abnormal in nonnumeric data (eg, flagging the presence of a physical sign as abnormal), it should consider age in the interpretation of normality. For example, if “unable to feed self” is considered to be a universally abnormal finding in the interpretation of a functional assessment, then the system is not taking the functional capabilities of young children into account.

**Complex Normative Relationships**
Not all normative data are based solely on age. In the case of blood pressure, normative values are determined by age (to the nearest month), gender, and height percentile. Similarly, peak flow meter norms depend on age, height, and gender. Methods for flagging abnormal values that are based on age alone are insufficient for blood pressure and peak expiratory flow and may be insufficient for other measurements in pediatric patients.

**Gestational Age**
For neonates, chronologic age (expressed simply as the time since birth) is insufficient for medication-prescribing decision support, normative ranges for laboratory data, normative definitions for physical examination findings, and guideline-application support. Gestational age, chronologic age, and corrected age are each unique and important ways to present age of a neonate; EHR systems need to record each of these expressions for age and allow for their use in decision support.

**Privacy**

**Adolescent Privacy**
Laws about age of consent vary from state to state and according to presenting problem. Adolescents who present for treatment of mental health disorders, for example, may consent to their treatment at an earlier age than the age of majority in most states. Some states also have laws regarding parental notification whereby interpretation is based on the patient’s age and presenting problem. Practices that serve adolescents typically have policies with respect to what portion of an adolescent’s care should be handled with special privacy protections (eg, in some jurisdictions, the adolescent must give explicit permission for the parent to review his or her records). These privacy protections may require the flagging of protected information. Therefore, EHR systems should support privacy policies that vary by age and according to presenting problem and diagnosis and be flexible enough to handle the policies of individual practices. Furthermore, if an EHR system handles record-keeping for consent for treatment, it should provide for the recording of assent for treatment (from an underaged adolescent or child) combined with parental informed permission as well as consent for treatment (from an adolescent) combined with a record of parental involvement. The separation of the patient’s consent and the parent’s or guardian’s consent is particularly important in the area of testing for drugs of abuse. Pregnancy is another area in which the records of patient and parental consent, assent, and permission may be less straightforward than in adult care.

**Children in Foster or Custodial Care**
When a child is removed from the care of his or her parents, as in the case of foster care, complex issues of confidentiality of medical information arise. Licensed foster parents may consent to routine medical and dental treatment for minors placed with them pursuant to a court order or with the voluntary consent of the person having the legal custody of the minor. The pediatrician should document the authority of a foster parent to give consent to medical treatment by obtaining a copy of the court order. Parents who no longer have custody may still have the right to access their children’s medical records and be involved with health care decisions unless their parental rights have been terminated. EHR systems that purport to manage consent for treatment and information access will need to be able to record these details.

**Consent by Proxy**
Children often present for nonurgent health care in the company of an adult who is not the custodial parent or
guardian. The best way to prevent confusion about consent for care in this situation is to record the custodial parents’ wishes as to which adult can consent to which child’s care and under what limitations. EHR systems that manage consent for treatment should support this sort of data-recording.

Adoption
Records of children who are undergoing adoption proceedings or who have been adopted may need special privacy handling, as in a case where state law offers special protections for the identity of adoptees. The EHR systems should allow flagging of these data for special privacy protection. In some states, the preadoption record may need to be separated entirely from any post-adoption record by using 2 distinct patient identities.

Guardianship
The identity of a child’s guardian and guarantor, although most commonly the parent, can become complicated outside the bounds of the “typical” 2-parent household. The EHR system must provide the flexibility to indicate the broad variety of adults in the child’s life who may play some role in medical or financial decision-making. The system should draw a distinction between the patient’s guardian and his or her financial guarantor. In those cases in which a court has appointed a guardian for a minor, the ability of the guardian to consent to medical treatment depends on the type of treatment being sought and the scope of authority the court has granted. If more than routine care is required, the pediatrician should document the authority of the guardian to give consent by obtaining a copy of the official certified letters of guardianship. The EHR system should support this record-keeping.

Emergency Treatment
When EHR systems support the recording of consent and assent for treatment, they should be flexible enough to allow for the emergency treatment of minors, in which the parent or legal guardian may be absent, and the usual procedures for consent must change.

PEDIATRIC TERMINOLOGY
Some of the barriers that child health care providers encounter in the application of EHR systems relate not to functions of the system but to the inappropriate terminology used to express concepts (eg, physical examination findings, developmental milestones, diagnoses) in the EHR system’s user interface. These terminology systems differ from systems such as the International Classification of Diseases, 9th Edition, Clinical Modification, which is used to encode diagnoses for insurance claims. Rather, these terminology systems are used to allow the precise encoding of clinical concepts by the user in lieu of free text. EHR systems generally use a terminology developed by a third party or by the EHR system developers themselves. A complete treatment of special terminology requirements is outside the scope of this report. The AAP and its members should advocate for the inclusion in these systems of historical findings, psychosocial risk factors, family structural details, social history, physical examination findings, developmental problems, behavioral issues, congenital syndromes, and diagnoses of particular importance to pediatrics. The US government’s Consolidated Health Informatics Initiative, which specifies which terminology system should be used in which clinical domain within government-sponsored health-information systems, should help focus the advocacy effort of the AAP. It is important to note, however, that no health-information system directly managed by the US federal government deals primarily with children.

DATA PRECISION
There is a broad category of functionality that may limit an EHR system’s usefulness in pediatric practice: the ability to handle data at an appropriate numeric precision and graphical resolution. For example, body weight to the nearest gram is commonly accepted as an appropriate precision in neonatal facilities. As another example, an EHR system may present growth curves of height, weight, and head circumference, complete with appropriate normative curves for comparison. However, if those curves are available in only 1 graphical resolution, measurements obtained frequently (daily weight measurements, weekly head circumference measurements, etc) may become impossible to analyze visually. Age in the newborn nursery should be expressed in units at least down to the hour, if not to the minute. The units for age (days, weeks, months, years) need to grow with the age of the child, as appropriate. Developers of EHR systems should consider how the small changes in numeric data that one sees in the care of young patients affect data-recording and display.

OTHER PEDIATRIC FUNCTIONS
This report outlines the major areas of functionality that are relatively more important in pediatric care than in adult care. There are, of course, many other functions that are important, such as the ability to:

- archive and manage patient data for a statutorily defined period of time;
- provide educational materials that are appropriate to both parents and children and at varying reading levels;
- create pedigree diagrams;
- display age at all times throughout the user interface;
- select age-based documentation templates and order sets on the basis of a patient’s age;
• indicate whether a guideline applies to a patient on the basis of age; and
• indicate the source of patient data, especially when the source is not the patient or the parent (eg, the school teacher or child care worker).

PEDIATRIC EHR SYSTEM FUNCTIONALITY STANDARDS

HL7 is an organization that was founded in 1987 to set international standards for how health information is exchanged between information systems. It expanded its scope beyond data interchange to include specifications for EHR system functions through its Electronic Health Record Technical Committee. The Electronic Health Record Technical Committee, which was founded in 2001, published its first balloted standard for EHR system functions in 2004. This standard is being used as the basis for the EHR system certification process specified by the federal Office of the National Coordinator for Health Information Technology (created by Executive Order 13335, April 28, 2004, and authorized by Congress [FR Doc No. 05-16446, Filed August 18, 2005]). The purpose of certification is to set a minimum level of functionality that EHR systems will have to meet to qualify for special treatment, such as participation in pay-for-performance programs. By contract with the Office of the National Coordinator for Health Information Technology, the Certification Commission for Health Information Technology (CCHIT [www.cchit.org]) is charged with establishing a certification process by which EHR system software may be declared eligible for pay-for-performance incentives designed to promote care facilitated by an information system. The CCHIT has several pediatricians working on its committees to ensure that pediatric functions are incorporated into the certification process. As of this writing, patient-care scenarios of the CCHIT that were designed to test functionality exclude infants. The HL7 Pediatric Data Standards Special Interest Group is working with the HL7 Electronic Health Record Technical Committee to ensure that the pediatric functions mentioned in this statement are included in the HL7 EHR functional model and, therefore, will become a part of EHR system certification processes in the future. The current EHR system functional model may be obtained from the HL7 Web site (www.hl7.org).

THE FUTURE OF THE PEDIATRIC EHR SYSTEM

In the wake of the rapid uptake of EHR systems in the years since the first AAP statement, national groups have expressed increased interest in standardizing the features of EHR systems and certifying their functions. Child health care providers want to be sure that pediatric functions, terminology, and data precision are built into these standards and certification processes. They want this not only to make their own systems more effective in improving the health of children but also to make all EHR systems more useful for patients of all ages. The AAP is working proactively to ensure that knowledgeable pediatricians who can thoroughly explain child health care issues are invited to address the groups that set these standards. This report should serve as a guide for these efforts to represent the interests of child health care providers and present a guide to individual practitioners who are evaluating a given system’s ability to perform in the pediatric environment.

COUNCIL ON CLINICAL INFORMATION TECHNOLOGY, 2006–2007

Mark M. Simonian, MD, Chairperson
Joseph H. Schneider, MD, MBA, Vice Chairperson
Kristin A. Benson, MD
Donna M. D’Alessandro, MD
Mark A. Del Beccaro, MD
Willa H. Drummond, MD, MS
George R. Kim, MD
Michael Leu, MD
Gregg C. Lund, DO
Eugenia Marcus, MD
Alan E. Zuckerman, MD

PAST COUNCIL MEMBERS

S. Andrew Spooner, MD, MS, Immediate Past Chairperson
Robert S. Gerstle, MD
Kevin B. Johnson, MD, MS
Christoph U. Lehmann, MD

STAFF

Rebecca Marshall

REFERENCES

38. Kemper AR, Uren RL, Clark SJ. Adoption of electronic health records in primary care pediatric practices. Pediatrics. 2006;118(1). Available at: www.pediatrics.org/cgi/content/full/118/1/e20
Self-injectable Epinephrine for First-Aid Management of Anaphylaxis

Scott H. Sicherer, MD, F. Estelle R. Simons, MD, and the Section on Allergy and Immunology

ABSTRACT
Anaphylaxis is a severe, potentially fatal systemic allergic reaction that is rapid in onset and may cause death. Epinephrine is the primary medical therapy, and it must be administered promptly. This clinical report focuses on practical issues concerning the administration of self-injectable epinephrine for first-aid treatment of anaphylaxis in the community. The recommended epinephrine dose for anaphylaxis in children, based primarily on anecdotal evidence, is 0.01 mg/kg, up to 0.30 mg. Intramuscular injection of epinephrine into the lateral thigh (vastus lateralis) is the preferred route for therapy in first-aid treatment. Epinephrine autoinjectors are currently available in only 2 fixed doses: 0.15 and 0.30 mg. On the basis of current, albeit limited, data, it seems reasonable to recommend autoinjectors with 0.15 mg of epinephrine for otherwise healthy young children who weigh 10 to 25 kg (22–55 lb) and autoinjectors with 0.30 mg of epinephrine for those who weigh approximately 25 kg (55 lb) or more; however, specific clinical circumstances must be considered in these decisions. This report also describes several quandaries in regard to management, including the selection of dose, indications for prescribing an autoinjector, and decisions regarding when to inject epinephrine. Effective care for individuals at risk of anaphylaxis requires a comprehensive management approach involving families, allergic children, schools, camps, and other youth organizations. Risk reduction entails confirmation of the trigger, discussion of avoidance of the relevant allergen, a written individualized emergency anaphylaxis action plan, and education of supervising adults with regard to recognition and treatment of anaphylaxis.

INTRODUCTION
Anaphylaxis is an acute, life-threatening reaction, usually mediated by an immunologic mechanism involving immunoglobulin E, that results in sudden systemic release of mast-cell and basophil mediators such as histamine and tryptase. Anaphylaxis has many clinical presentations, but respiratory compromise and cardiovascular collapse cause the greatest concern, because they can potentially lead to fatalities. Although a variety of different triggers for anaphylaxis episodes have been identified, food and insect stings are the most common identifiable triggers reported in the community setting. Food allergies and other allergies have increased in the past several years, and pediatricians increasingly need to prescribe emergency care plans for patients in the event of anaphylaxis outside the hospital/medical setting. Epinephrine is the primary medical therapy for a life-threatening allergic reaction. This clinical report focuses on practical issues concerning the administration of self-injectable epinephrine for first-aid treatment of anaphylaxis.

638 AMERICAN ACADEMY OF PEDIATRICS
anaphylaxis in the community. In addition, several quandaries in management will be identified and possible solutions described.

**DEFINITION AND FEATURES OF ANAPHYLAXIS**

There is no current, universally accepted definition of anaphylaxis; however, at a recent symposium cosponsored by the National Institutes of Health and the Food Allergy & Anaphylaxis Network, the following definition was proposed: “Anaphylaxis is a serious allergic reaction that is rapid in onset and may cause death.”6,7 Three clinical criteria for anaphylaxis based on symptoms and history were also proposed at the symposium. These criteria, as well as various signs and symptoms that may occur during anaphylaxis, are listed in Table 1. The clinician must also appreciate that certain disorders may appear to be anaphylaxis but are not (eg, vasovagal syncope or panic attack). Fatal anaphylaxis in the pediatric population has particularly been associated with known preexisting asthma, failure to administer epinephrine promptly, and the adolescent age group.8,9

### TABLE 1 Clinical Criteria for Diagnosing Anaphylaxis (Fulfilling Any 1 Criterion Indicates That Anaphylaxis Is Highly Likely)

<table>
<thead>
<tr>
<th>Criterion 1</th>
<th>Acute onset of an illness (minutes to several hours) with involvement of the skin and/or mucosal tissue (eg, generalized hives, pruritus, or flushing, swollen lips/tongue/uvula) and at least 1 of the following: a. Respiratory compromise (eg, dyspnea, wheeze/bronchospasm, stridor, reduced peak expiratory flow, hypoxemia) b. Reduced blood pressure or associated symptoms of end-organ dysfunction (eg, hypotonia [collapse], syncope, incontinence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion 2</td>
<td>Two or more of the following that occur rapidly after exposure to a likely allergen for that patient (minutes to several hours): a. Involvement of the skin/mucosal tissue (eg, generalized hives, itch/flush, swollen lips/tongue/uvula) b. Respiratory compromise (eg, dyspnea, wheeze/bronchospasm, stridor, reduced peak expiratory flow, hypoxemia) c. Reduced blood pressure or associated symptoms (eg, hypotonia [collapse], syncope, incontinence) d. Persistent gastrointestinal symptoms (eg, crampy abdominal pain, vomiting)</td>
</tr>
<tr>
<td>Criterion 3</td>
<td>Reduced blood pressure after exposure to known allergen for that patient (minutes to several hours)</td>
</tr>
</tbody>
</table>

Less common presentations also occur (eg, sudden isolated hypotension without a known allergen exposure). Additional symptoms and signs that may occur during anaphylaxis include mobiliform rash, conjunctival erythema, pruritus and tightness in the throat, dysphagia, dysphonias, dry staccato cough, sensation of pruritus in the external auditory canals, nasal pruritus, nasal congestion, rhinorrhea, sneezing, chest pain, dysrhythmia, feeling of faintness/dizziness (near-syncope), pallor, cyanosis, confusion/altared mental status, an aura of doom, and atetonic contractions. Skin signs aid in recognition but may be absent or not observed in 10% of children with anaphylaxis; moreover, they may not be observed in reactions that end in fatality.


### ROLE OF EPINEPHRINE

Epinephrine, the medication of choice for first-aid treatment of an episode of anaphylaxis, is a direct-acting sympathomimetic agent with effects on many target organs, including increased vasoconstriction, decreased mucusal edema, increased inotropy/chronotropy, and bronchodilation. In addition, epinephrine downregulates further mast-cell release of histamine, tryptase, and other mediators of inflammation. Delayed administration of epinephrine in anaphylaxis is associated with poor outcomes including fatality.6–11 Oral H1 antihistamines are not an optimal first-line therapy for anaphylaxis, because they have a slow onset of action (1 or more hours), primarily relieve cutaneous symptoms, and do not relieve respiratory symptoms or shock.12 For children with known preexisting asthma who experience anaphylaxis, administration of an asthma-reliever medication (such as the inhaled selective β2-adrenergic agonist albuterol) may provide adjunctive therapy for wheezing, coughing, and shortness of breath but does not relieve upper airway edema or shock, and therefore does not replace injected epinephrine in anaphylaxis management.

### ROUTE OF ADMINISTRATION OF EPINEPHRINE

For first-aid treatment of anaphylaxis, administration of epinephrine by either the subcutaneous or intramuscular route has been recommended traditionally. However, studies on the rate of absorption of epinephrine injected by different routes and in different locations (eg, arm or thigh) have shown significant differences in time to peak concentrations, favoring intramuscular injection in the lateral thigh (vastus lateralis muscle), which leads promptly to peak plasma epinephrine concentrations. In a prospective, randomized, blinded study of children at risk of anaphylaxis,13 the time to maximum epinephrine concentrations was 8 ± 2 minutes after injection of 0.30 mg of epinephrine from an EpiPen (Dey LP, Napa, CA) intramuscularly in the vastus lateralis. In contrast, the time to maximum plasma epinephrine concentration was 34 ± 14 minutes (range: 5–120 minutes) after injection of 0.01 mg/kg of epinephrine subcutaneously in the deltoid region.11 These findings have been confirmed and extended in a randomized, double-blind, placebo-controlled crossover study in adults.14 On the basis of studies in an animal model of anaphylaxis, achieving high plasma and tissue concentrations of epinephrine may be critical for reversal of hypotension.15 The ½-inch (14.29-mm) needle on autoinjectors likely provides an intramuscular dose in most children, although it may not do so in obese adolescents, especially girls.16

It is not ethical to perform randomized, double-blind, placebo-controlled comparative studies on route of administration of epinephrine in children who are experiencing anaphylaxis, so definitive evidence-based recommendations on route of dosing cannot be made. On the
basis of the aforementioned available data at this time, intramuscular injection of epinephrine into the lateral thigh seems to be the preferred route for therapy in first-aid treatment. Intravenous administration carries risks of dilution errors and dosing errors, and many of the serious adverse effects attributed to epinephrine have followed large overdoses given intravenously. This route of administration should be reserved for those with severe anaphylaxis that does not respond to intramuscular epinephrine and/or individuals with anaphylaxis who are being treated in hospital settings.

**EPINEPHRINE DOSING FOR FIRST-AID TREATMENT IN THE COMMUNITY**

The recommendation for epinephrine dosing in children with anaphylaxis, based primarily on anecdotal evidence, is to inject 0.01 mg/kg, up to 0.30 mg.17–19 Epinephrine autoinjectors are currently available in 2 fixed doses: 0.15 and 0.30 mg. Physicians, therefore, face a quandary with regard to dosing children who do not weigh approximately 15 kg (33 lbs [for whom the 0.15-mg dose is ideal]) or 30 kg (66 pounds or more [for whom the 0.30-mg dose is recommended]). The Physician’s Desk Reference20 and product inserts provide ambiguous advice and place the responsibility of dose selection entirely on the prescribing physician. Not surprisingly, therefore, both autoinjector doses are dispensed across almost the entire pediatric age range, indicating the potential for overdosing with the 0.15-mg dose in many infants, overdosing with the 0.30-mg dose in some young children, and underdosing with the 0.15-mg dose in many adolescents.21 In a prospective, randomized, double-blind, parallel-group study of children at risk of anaphylaxis who self-injected either EpiPen 0.30 mg or EpiPen Jr 0.15 mg, pharmacologic effects such as pallor, tremor, and anxiety were observed transiently after injection of both doses, and additional effects including palpitations, headache, and nausea were observed in those who received the higher dose.22 This study showed that the therapeutic effects of epinephrine could not be dissociated from the nontherapeutic effects. In the absence of availability of additional fixed doses (eg, 0.05, 0.10, 0.20, and 0.25 mg) in autoinjectors, the manufacturers’ advice should be taken in the context of these benefits and risks.

**The Risks of Prescribing an Epinephrine Ampule, Syringe, and Needle**

Physicians face a particularly difficult dilemma in prescribing epinephrine doses for infants and children who weigh less than 15 kg (33 lb). One option may be to prescribe an epinephrine ampule/syringe/needle and instruct caregivers on how to draw up and inject epinephrine using these supplies. This approach was studied in 18 parents who were trained in the technique and whose speed and accuracy of drawing up an infant epinephrine dose (0.09 mL) was compared with that of 54 physicians and nurses (controls).23 The parents took significantly (P < .05) longer than the controls to draw up the dose. The mean ± SEM times for drawing up doses were 142 ± 13 seconds (range: 83–248 seconds) for parents, 52 ± 3 seconds (range: 30–83 seconds) for physicians, 40 ± 2 seconds (range: 26–71 seconds) for general duty nurses, and 29 ± 0.09 seconds (range: 27–33 seconds) for emergency department nurses. The epinephrine content of the doses drawn up by parents, who were asked to draw up 0.09 mL, ranged from 0.004 to 0.151 mL (ie, nearly 40-fold). There was no correlation between speed of drawing up the epinephrine and accuracy of dosing. Parents had many concerns about successfully preparing and administering a dose by this method and about teaching other caregivers to use the method. The study was undertaken in a relaxed atmosphere, and one might expect more difficulties if the dosing were undertaken by laypersons while a child was experiencing anaphylaxis.24

Unfortunately, the current lack of autoinjectors with a low or adjustable dose and the problems involved with prescribing an epinephrine ampule along with a syringe/needle require that the physician and family arrive at a reasonable compromise for safe and effective administration of epinephrine in the event of anaphylaxis. In a survey of 29 pediatricians, 80% responded that they would prescribe the 0.15-mg autoinjector dose for a child who weighs 10 kg (22 lb); 100% responded that they would prescribe it for a child who weighs 15 kg (33 lb); and 70% responded that they would prescribe it for a child who weighs 20 kg (44 lb).25 In a study of epinephrine-dispensing patterns,21 72% of prescriptions for infants younger than 6 months (weighing less than approximately 7 kg [15 lb]) were for a 0.15-mg autoinjector, and 20% were for ampule/syringe/needle; 95% of prescriptions for infants 6 to 12 months of age (likely weighing up to approximately 10 kg [22 lb]) were for a 0.15-mg autoinjector. Until a wider range of epinephrine autoinjector doses is available, pediatricians are forced to consider prescribing an autoinjector with a known, albeit not ideal, dose rather than risk likely overdosing or underdosing with ampule/syringe/needle.

**Epinephrine Autoinjectors: 0.15 or 0.30 mg?**

In the absence of a strong evidence base and a larger selection of premeasured autoinjector doses, and in light of studies showing elevated plasma concentrations and relatively modest adverse effects in children who weigh approximately 25 kg (55 lb) injected with approximately a 1.2-fold overdose of intramuscular epinephrine, it seems appropriate to switch most children from the 0.15-mg dose to the 0.30-mg dose at approximately 25 kg (55 lb)—that is, to provide a slightly higher dose (0.012 mg/kg) rather than an underdose (0.06 mg/kg) for a 25-kg (55-lb) child. For children who have asthma
or other additional risk factors for fatality from anaphylaxis, switching to the higher dose at a lower weight might be considered. There are no data at this time to support specific recommendations for children who weigh less than 15 kg (33 lb). It is not known whether the adverse effects from a previously studied 1.2-fold overdose are similar in infants and very young, small children to those reported in children who weigh approximately 25 kg (55 lb). Considering the ease of use of self-injectable epinephrine compared with the ampule/syringe/needle technique and the preferences of pediatricians and parents, it seems reasonable to consider autoinjectors containing 0.15 mg of epinephrine for otherwise healthy infants/young children who weigh 10 to 25 kg (22–55 lb), although the physician is cautioned that manufacturers suggest alternative modalities for those who weigh less than 15 kg (33 lb). Clearly, specific clinical circumstances must be considered in this decision. For infants who weigh less than 10 kg (22 lb), dosing with 0.15-mg autoinjectors would exceed 1.5-fold overdosage, and although this situation is unacceptable from the standpoint of autoinjector availability, it is apparent that many pediatricians opt for the certainty of an autoinjector dose compared with the uncertainty of an ideal dose when the epinephrine ampule/syringe/needle technique is used. Still, physicians and families should consider and discuss the benefits and risks of choosing between an autoinjector or epinephrine ampule/syringe/needle for this age group on a case-by-case basis. Current dilemmas in selecting a dose and factors that may sway a decision to switch to a particular dose unit are listed in Table 2. Lack of worldwide availability of autoinjectors often requires prescription of the less costly epinephrine ampule/syringe/needle technique in developing countries despite the fact that it requires additional training, is error prone, and may lead to delay in injection. Preloading the syringe with an appropriate dose of epinephrine is a possible partial solution, but contamination and degradation of the drug, particularly in hot climates, are serious concerns.

### Repeating the Epinephrine Dose
Anecdotal evidence generally suggests that in the absence of a response to epinephrine, the epinephrine injection may be repeated at 5- to 20-minute intervals. Retrospective studies have suggested that a sec-

<table>
<thead>
<tr>
<th>Patient’s Weight, kg (lb)</th>
<th>Optimal Dose (0.01 mg/kg), mg</th>
<th>Availability of Autoinjector</th>
<th>Alternatives/Implications</th>
<th>Comments/Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤10 (≤22)</td>
<td>≤0.10</td>
<td>No</td>
<td>Fixed-dose 0.15-mg autoinjector provides 1.5-fold overdose; ampule/syringe/needle technique may lead to delay in injection and inaccurate dosing</td>
<td>Evaluate degree of overdose vs ability to use ampule/syringe/needle; no specific evidence base for decision except that ampule/syringe/needle technique is delay and error prone, and autoinjector (0.15 mg) is more commonly prescribed for infants by physicians forced to choose</td>
</tr>
<tr>
<td>15 (33)</td>
<td>0.15</td>
<td>Yes</td>
<td>0.15-mg autoinjector provides optimum dose</td>
<td>Prescribe autoinjector (0.15 mg)</td>
</tr>
<tr>
<td>20 (44)</td>
<td>0.20</td>
<td>No</td>
<td>0.15-mg autoinjector provides 1.3-fold overdose; 0.30-mg autoinjector provides 1.5-fold overdose</td>
<td>Usually prescribe 0.15-mg autoinjector, but increasing weight of child over 20 kg and high risk on the basis of clinical history may be considered an appropriate rationale for prescribing a 0.30-mg autoinjector</td>
</tr>
<tr>
<td>25 (55)</td>
<td>0.25</td>
<td>No</td>
<td>0.15-mg autoinjector provides 1.7-fold overdose; 0.30-mg autoinjector provides 1.2-fold overdose</td>
<td>Usually prescribe 0.30-mg autoinjector, a small overdose in a healthy child generally carries a low risk of adverse effects compared with the risk of an underdose during anaphylaxis</td>
</tr>
</tbody>
</table>

* For situations in which an autoinjector containing an appropriate dose is not available, the situation is never truly acceptable, because using an epinephrine ampule/syringe/needle (see text) is prone to delay in dosing or inaccurate dosing. However, until such autoinjectors are manufactured and fixed doses of 0.05, 0.10, 0.20, and 0.25 mg are available in addition to the 0.15- and 0.30-mg doses currently available, the physician has to determine the risk versus the benefit of selecting a fixed dose that is either too low or too high and the risk/benefit of an optimal technique of administration (autoinjector) versus a technique (ampule/syringe/needle) that may be prone to delay and error in the hands of non–health care professionals.
* There are no studies that have provided details about risk of overdose and underdose of epinephrine in the context of first aid treatment of anaphylaxis at most dose ranges, particularly in children who weigh less than 15 kg (33 lb). It is presumed, on the basis of limited data, that otherwise healthy children (normal cardiac status, not taking other sympathomimetics, tricyclic antidepressants, or monoamine oxidase inhibitors, etc.) would tolerate modest overdoses of epinephrine. In older children not experiencing anaphylaxis, a 1.2-fold overdose has been associated with adverse pharmacologic effects.
* Distributors’ recommendations regarding autoinjector indications for weight/age differ from country to country, but an alternative form of epinephrine for self-injection, such as ampule/syringe/needle rather than an autoinjector, has been suggested for children who weigh less than 15 kg (33 lb). The perceived comfort and ability of families and caregivers to provide accurate doses of epinephrine for infants using an epinephrine ampule/syringe/needle should be considered in deciding the best modality and the potential degree of overdose or underdose if an autoinjector were prescribed. Although not per manufacturer’s advice, it is suggested that the available evidence (error rates of ampule/syringe/needle of no dose to almost 40-fold overdose in the hands of non–health care professionals, adverse pharmacologic effects of a modest overdose in a healthy child, lack of additional fixed-dose autoinjectors) may warrant prescription of a 0.15-mg autoinjector for most healthy children who weigh 10 kg (22 lb) or more; however, individual circumstances may vary. Depending on the circumstances, provision of an autoinjector to those who weigh less than 10 kg (22 lb) may also be warranted.
* In addition to consideration of body weight, clinical issues that may add risk to underdosing and indicate a relative benefit for a higher dose may include 1 or more of the following: concurrent asthma; previous anaphylaxis to peanut, tree nut, milk, egg, seafood, and/or fin fish; poor access to emergency services; and/or lack of supervision.
ond dose may be required in 18% to 35% of cases, although data in this regard are limited.29,30 As stated previously, some of the effects of epinephrine (pallor, tremor, anxiety, and palpitations) and even severe adverse effects (such as cough from pulmonary edema) can mimic some of the symptoms of anaphylaxis. Caregivers should be aware of these issues and avoid unnecessary repeat dosing.

In some adults experiencing anaphylaxis who were raised from the supine to the upright position during transport to a hospital, death occurred suddenly, presumably from an “empty-ventricle syndrome” caused by blood pooling in the legs during anaphylactic shock.31 The implications of this observation for children, who more typically succumb to respiratory insufficiency during anaphylaxis and who often vomit during anaphylaxis, are not known. Nevertheless, caregivers should be advised that individuals with severe anaphylaxis who may benefit from being in a supine position with legs raised should remain in that position and be transported that way by emergency personnel until advanced care can be accessed (eg, additional medications and intravenous fluids).

**PRESCRIPTION OF SELF-INJECTABLE EPINEPHRINE**

The primary indication for prescription of self-injectable epinephrine is a history of anaphylaxis in an individual who may re-encounter the triggering agent outside of a medical setting or who has idiopathic anaphylaxis, which is uncommon in childhood. Identification of individuals who have experienced anaphylaxis is not necessarily easy. It is clear that persons with a previous episode of anaphylaxis that was characterized by respiratory or cardiovascular compromise to a trigger that may be encountered outside the hospital should carry self-injectable epinephrine, but only approximately 70% of individuals with anaphylaxis have respiratory symptoms, and even fewer (only approximately 10%) experience cardiovascular symptoms.32 Skin manifestations such as urticaria, angioedema, flushing, or itching occur in more than 80% of children with anaphylaxis. When present, these symptoms are helpful in the recognition of anaphylaxis; when absent, they make the recognition of anaphylaxis more difficult.32 Moreover, acute generalized urticaria and angioedema alone may not necessarily warrant a diagnosis of “anaphylaxis” (a point of controversy). However, on the basis of available evidence, self-injectable epinephrine should be prescribed for a child who has experienced generalized acute urticaria after an insect sting, because the risk of a more severe reaction from a future sting is approximately 10%.33 Finally, physicians cannot assume that patients and caregivers necessarily recognize and report all symptoms, because even trained health care professionals underrecognize anaphylaxis.3 For all of these reasons, a high index of suspicion is needed to identify those who have had anaphylaxis and require an epinephrine prescription.

An additional point of judgment regarding prescription of self-injectable epinephrine is that a physician may identify a child who has not yet experienced anaphylaxis but may nevertheless be at increased risk of anaphylaxis and may warrant prescription of self-injectable epinephrine. Vander Leek et al34 showed that among 24 young children with peanut allergy whose first reaction was isolated to the skin after ingestion or skin contact, 18 (75%) experienced symptoms beyond the skin in a subsequent reaction. Indeed, severity of a previous reaction is a poor guide to symptoms during a future reaction.9,11,35 In young children with peanut- or tree nut–related anaphylaxis, episodes may worsen progressively with time, perhaps related to the fact that increased numbers of such children develop asthma as they get older.26 Asthma, which is associated with severe and fatal anaphylaxis,8,9,28 is an important comorbidity that should influence the decision to prescribe self-injectable epinephrine. Some “high-risk” circumstances that may justify prescription of self-injectable epinephrine in the absence of previous anaphylaxis are summarized in Table 3.18,37,38 Definitive evaluations of such children by an allergy/immunology specialist with American Board of Allergy and Immunology certification or international equivalent should be encouraged.39

In summary, epinephrine should be prescribed for children who have experienced anaphylaxis and may re-encounter the trigger outside of a hospital setting. In some circumstances, epinephrine for self-injection should be prescribed for persons who have not experienced anaphylaxis but are at increased risk of anaphylaxis on the basis of their specific comorbid medical conditions and medical-social evaluation.

**INSTRUCTIONS FOR WHEN TO USE EPINEPHRINE**

Physicians should carefully instruct patients and families on the indications for, and the technique for using,
self-injectable epinephrine. Prompt administration of epinephrine is clearly indicated for treatment of significant respiratory or cardiovascular symptoms of anaphylaxis, but considerable judgment is required in many actual or possible allergic reactions in which life-threatening symptoms have not yet developed but may develop. Previous guidelines have suggested that epinephrine should be administered promptly at the onset of symptoms after exposure to an allergen that had previously caused anaphylaxis and possibly even in the absence of symptoms if there was a known exposure to an allergen that previously caused anaphylaxis with cardiovascular collapse. Generalized acute urticaria itself is not a life-threatening symptom, yet in the context of a known exposure to an allergen that previously triggered anaphylaxis, the recommendation for an exposure outside of a medical setting is to inject epinephrine. Whether an individual with generalized acute urticaria has “anaphylaxis” and should be given epinephrine is controversial. In many circumstances, astute clinical judgment is required to differentiate symptoms that may mimic aspects of an episode of anaphylaxis (eg, viral syndrome with acute urticaria, asthma, choking, a panic episode) or represent a mild allergic reaction that does not require epinephrine. In the community setting, individuals who experience anaphylaxis, whose judgment may be clouded by anxiety or central nervous system symptoms, or caregivers without medical training, whose judgment may be clouded by anxiety, are required to evaluate symptoms. Consequently, physicians should always instruct these individuals to err on the side of injecting epinephrine rather than waiting too long.

Individuals and caregivers are often reluctant to use self-injectable epinephrine in anaphylaxis despite instruction to do so. This probably occurs for a variety of reasons, including failure to recognize anaphylaxis; spontaneous recovery from a previous episode; incorrectly thinking the episode is mild; reliance on oral H1 antihistamines or asthma-relief inhalers such as albuterol; fear of needles and injections; epinephrine auto-injector not being available; and concern about adverse effects of epinephrine. In contrast to transient pallor, tremor, anxiety, and palpitations, which are common and anticipated pharmacologic effects of epinephrine, serious adverse effects are generally not a concern for otherwise healthy children, although they have been reported when epinephrine was given in overdose, especially when it was administered intravenously in an overdose, given at an inappropriately high concentration, or infused too rapidly.

It seems that adolescents are at particular risk of fatal anaphylaxis, possibly because they are more likely to engage in risky behaviors, fail to recognize triggers, deny symptoms, and not carry or use emergency medications. Additional efforts to provide anaphylaxis education for adolescents and their friends and peers are needed.

Prompt administration of epinephrine for anaphylaxis is key. Sampson et al described 6 children with fatal reactions to food, all of whom had asthma, previous reactions to foods, and delay in treatment with epinephrine. None of the children received epinephrine before onset of severe respiratory symptoms (obvious respiratory distress, retractions, wheezing, and, in some cases, cyanosis), and 7 children in the same study with near-fatal food anaphylaxis received epinephrine before or within 5 minutes of severe respiratory symptoms. Only 1 of the children with fatal reactions had cutaneous symptoms; in contrast, all of those with near-fatal reactions had cutaneous symptoms. This raises the concern that absence of, or failure to recognize, skin symptoms and other symptoms could result in a delay in treatment and a poor outcome. Among 32 food-anaphylaxis fatalities recorded in a registry maintained through the Food Allergy & Anaphylaxis Network, all but 1 individual had a known allergy to the food, only 10% had self-injectable epinephrine available, peanut or tree nut caused 94% of the reactions (milk and fish caused the others), most of those who died were adolescents or young adults, and 96% had asthma.

Gold and Sainsbury surveyed families of children for whom self-injectable epinephrine was prescribed for a previous reaction with respiratory or cardiovascular involvement. Although recurrences were common, epinephrine was injected in only 12% of subsequent reactions. When it was given, although it was seldom injected before onset of respiratory or cardiovascular symptoms, it resulted in a significantly lower hospitalization rate and reduced morbidity.

When developing an anaphylaxis emergency action plan for an individual to use in the community in the absence of a health care professional, presumably for a circumstance in which definitive diagnosis is unlikely, it seems advisable to instruct patients/caregivers to inject epinephrine promptly when symptoms occur after known exposure to a trigger that previously caused a significant reaction. For the occasional child or adolescent who has idiopathic anaphylaxis, where “exposure” is an irrelevant issue, a symptom-based approach is required.

Patients and caregivers must also be instructed in the techniques of autoinjector use or epinephrine ampule/syringe/needle use. Although the autoinjector devices are not particularly difficult to use, errors are common. The injection may be given through clothing, although care must be taken to avoid obstructions such as seams or items in pockets. Accidental injection of epinephrine into a digit can cause vasoconstriction and necrosis and should be promptly evaluated and treated, if necessary, with warming, topical nitroglycerin cream, or locally injected phenolamine or other vasodilator.
and practice of injection technique using “trainers” and review of manufacturer’s educational materials (eg, DVDs) are strongly recommended. Proper storage of the epinephrine, away from extremes of temperature and direct sunlight to protect the drug from degradation, is also important. Degradation may occur without discoloration or precipitation.31 It is important to remind patients and families to check autoinjector expiration dates and renew prescriptions promptly.

Preparation for first-aid treatment of anaphylaxis additionally requires medical home development and review of a personalized anaphylaxis emergency action plan that lists potential anaphylaxis symptoms and gives instructions for the indications for self-injectable epinephrine, the technique for using epinephrine autoinjectors, and the necessity of taking the patient to an emergency department after an epinephrine injection. Downloadable examples of written plans that can be personalized are available through the Food Allergy & Anaphylaxis Network Web site (www.foodallergy.org/actionplan.pdf) [in English] or www.foodallergy.org/spanishaction.pdf [in Spanish]) and from the American Academy of Allergy, Asthma and Immunology Web site (www.aaaai.org/members/resources/anaphylaxis_toolkit/actionplan.pdf).39 The emergency action plan and coaching with regard to use of self-injectable epinephrine should be reviewed with the patient on a regular basis. Additional important considerations include diagnostic confirmation/reconfirmation of the triggering allergen, instructions with regard to trigger avoidance (for foods, insect stings, etc), and medical identification (eg, bracelet, wallet card).9 When relevant, specific preventive measures should be recommended (eg, for venom anaphylaxis, allergen-specific immunotherapy should be instituted to provide long-lasting protection).44 For exercise-induced anaphylaxis, physicians should recommend appropriate avoidance of food or medication co-triggers, and if no co-trigger has been identified, they should advise individuals to avoid ingestion of anything within 3 to 4 hours of strenuous exercise. Evaluation by an allergy/immunology specialist (with American Board of Allergy and Immunology or international equivalent certification) is typically required to address these issues. Lay organizations such as the Food Allergy & Anaphylaxis Network (www.foodallergy.org) are an important resource for educational materials and support. Omission of these preventive strategies may contribute to poor outcomes.9,11

In summary, epinephrine is the drug of choice for first-aid treatment of anaphylaxis and should be injected promptly in the event of an anaphylactic reaction or when progression to anaphylaxis is likely and advanced care is not promptly available. Asthma puffers and/or antihistamines cannot be depended on in anaphylaxis.39

SPECIAL ISSUES FOR SCHOOLS

Protection of children at risk of anaphylaxis while in school, child care, or camp requires a concerted effort.28 Several organizations have developed thoughtful summaries of shared responsibilities concerning food allergies for use by schools, children, adolescents, and parents (a list is available online at www.foodallergy.org/school/SchoolGuidelines.pdf). The physician should work with school administrators, teachers, school nurses, and others to ensure that an appropriate diagnosis has been obtained and that an appropriate anaphylaxis emergency action plan is prescribed.

SUMMARY

1. Epinephrine is the medication of choice for first-aid treatment of an episode of anaphylaxis. Prompt injection of epinephrine is nearly always effective in the treatment of anaphylaxis, and delayed injection of epinephrine is associated with poor outcomes including fatality. Antihistamines and, for those with asthma, inhaled selective β2-adrenergic agonists such as albuterol provide adjunctive therapy but cannot replace epinephrine. Advanced care for anaphylaxis should be sought promptly (call 911 or equivalent for additional care and emergency transport to a hospital/emergency department) after epinephrine injection for first-aid treatment of anaphylaxis.

2. The recommended epinephrine dose for anaphylaxis in children, based primarily on anecdotal evidence, is 0.01 mg/kg, up to 0.30 mg.

3. On the basis of the available data at this time, intramuscular injection of epinephrine into the lateral thigh (vastus lateralis) seems to be the preferred route for therapy in first-aid treatment, assuming that an early peak epinephrine concentration is important to effective management. Intravenous administration of epinephrine carries increased risks of dilution errors and dosing errors, with consequent increased risk of overdose and adverse effects such as cardiac dysrhythmias.

4. Epinephrine autoinjectors, preferred for ease of use compared with an ampule, syringe, and needle, are currently available in only 2 fixed doses: 0.15 and 0.30 mg. The lack of additional autoinjector doses is a serious concern. Nevertheless, pediatricians are advised to prescribe the optimal dose from an autoinjector for each child, even when that dose cannot possibly be precisely 0.01 mg/kg. On the basis of current, albeit limited, data, it seems reasonable to recommend autoinjectors with 0.15 mg of epinephrine for otherwise healthy young children who weigh 10 to 25 kg (22–55 lb) and autoinjectors with 0.30 mg of epinephrine for those who weigh approximately 25 kg (55 lb) and more. However, specific clinical
circumstances must be considered when making these decisions. For children who weigh less than 10 kg (22 lb), the physician and family should weigh the risks of delay in dosing and dosing errors when an ampule/syringe/needle is used against accepting non-ideal autoinjector doses, taking into consideration the specific health needs of the individual child and abilities of the caregivers.

5. Epinephrine should be prescribed for children who have experienced anaphylaxis who may re-encounter the trigger outside of a health care setting. In some circumstances, epinephrine for self-injection should be prescribed for persons who have not yet experienced anaphylaxis but are at increased risk of anaphylaxis on the basis of their specific trigger for anaphylaxis, comorbid medical conditions such as asthma, and/or limited ability to recognize anaphylaxis.

6. Epinephrine should always be prescribed in the context of an anaphylaxis emergency action plan developed by the medical home with the families. Effective care for individuals at risk of anaphylaxis requires a comprehensive management approach. Patients and caregivers must be carefully instructed on the technique for use of, and indications for, self-injectable epinephrine, how to recognize the symptoms of anaphylaxis, and the need to activate emergency services (call 911 or equivalent) in the event of anaphylaxis. Instructions on allergen avoidance are key. Optimal evaluation by an allergy/immunology specialist with American Board of Allergy and Immunology or international equivalent certification should be obtained to confirm allergic triggers, to provide education on trigger avoidance, and to initiate specific preventivetreatment (eg, venom-injection immunotherapy for insect-sting anaphylaxis). Written emergency action plans and review of care plans in the child’s medical home with specific responsibilities for school, child care, or camp personnel; families; and children are needed to ensure a safe environment for those at risk.

REFERENCES

1. American Academy of Allergy, Asthma and Immunology; American College of Allergy, Asthma and Immunology; Joint Council of Allergy, Asthma and Immunology; Joint Task Force on Practice Parameters. The diagnosis and management of anaphylaxis: an updated practice parameter. J Allergy Clin Immunol. 2005;115(3 suppl 2):S483–S523


Pulmonary Manifestations in ΔF508/R117H

To the Editor.—

O’Sullivan et al reported that children with the ΔF508/R117H-7T genotype may have pulmonary manifestations of cystic fibrosis (CF). CF lung disease is characterized by chronic endobronchial infection and inflammation, which leads to progressive airflow obstruction. Here I present evidence that infants carrying this genotype can have airflow obstruction even in the absence of pulmonary symptoms.

The state of New York initiated CF newborn screening in October 2002 with a system similar to that of Massachusetts. C. K. was a term product of an uncomplicated pregnancy, labor, and delivery. His immunoreactive trypsinogen was elevated, which prompted CF transmembrane conductance regulator (CFTR) gene mutation analysis. This showed him to be a compound heterozygote for ΔF508 and R117H. Intron analysis showed 1 copy each of 7T and 9T. His initial sweat chloride level was of insufficient quantity for analysis, but a repeat test at 6 weeks of age revealed a chloride level of 40 mmol/L. Results of an oropharyngeal swab obtained at 5 months of age revealed Pseudomonas aeruginosa. He received treatment with inhaled tobramycin (300 mg twice daily) for 4 weeks. Subsequent culture results obtained at quarterly intervals have been positive for Haemophilus influenzae and Staphylococcus aureus. The infant is now 17 months old, and he has had no radiographic, physical examination, or parent-reported signs or symptoms of lower respiratory tract disease. However, infant pulmonary-function tests performed at 15 months of age using raised-volume/rapid thoracoabdominal compression and body plethysmography showed evidence of obstructive lung disease, with a diminished forced expiratory flow at 75% of forced vital capacity of 46% predicted and an elevated functional residual capacity of 133% predicted.

This patient and all the patients reported by O’Sullivan et al had sweat chloride levels outside the 95% confidence interval for normal newborns, which suggests that in infants with this genotype, even sweat chloride values of 20 to 30 mmol/L do not rule out the possibility of pulmonary involvement. This case demonstrates that patients with the ΔF508/R117H-7T genotype can have airway obstruction even in the absence of demonstrable radiographic or clinical evidence of lung disease. These results support the 15 conclusions of O’Sullivan et al that patients with this genotype warrant close observation and follow-up. Infant pulmonary-function tests may be helpful in identifying airflow obstruction, which would support the diagnosis of CF in these infants.

Clement L. Ren, MD
Department of Pediatrics
University of Rochester
Rochester, NY 14642

REFERENCES

In Reply.—

We appreciate Dr Ren’s comments regarding our report of 4 children with the genotype ΔF508/R117H-7T. He
reports the case of another child identified by a positive immunoreactive trypsinogen/DNA cystic fibrosis (CF) newborn screen with a ΔF508/R117H-7T genotype who grew Pseudomonas aeruginosa from an oropharyngeal culture and demonstrated airway obstruction on infant pulmonary-function testing. We believe these 5 cases, along with those reported by Lording et al,2 supply ample evidence that the ΔF508/R117H-7T genotype is not necessarily benign. This experience contrasts with the conclusions of the report by Scotet et al3 in the November issue of Pediatrics. They reported a series of infants with elevated immunoreactive trypsinogen and the ΔF508/R117H-7T genotype and concluded that there was significantly better pulmonary function in the R117H compound heterozygotes as compared with age-matched controls with 2 severe CF mutations. Unfortunately, FEV1 (forced expiratory volume in 1 second) data were only available for 2 of 9 R117H compound heterozygotes. Also, although only 1 of 9 R117H compound heterozygotes versus 5 of 9 controls grew P aeruginosa from respiratory cultures, the authors did not state how many of the R117H compound heterozygotes had sputum or oropharyngeal cultures obtained or how frequently they were obtained.

Scotet et al3 recommend removal of R117H from the CF newborn screening mutation panel on the basis of this very limited data set. Our findings,1 as well as those of Ren and Lording et al,2 demonstrate that children with this mutation are at some risk of Pseudomonas colonization and infection, respiratory symptoms, and abnormal pulmonary-function test results. We believe it is premature to recommend removal of R117H from CF newborn screening mutation panels. Larger longitudinal studies need to be performed. We continue to recommend that all infants identified with R117H-7T as 1 of 2 CF transmembrane conductance regulator (CFTR) mutations be followed by a CF clinician, including evaluation for CF flora, chest imaging, and pulmonary-function testing, when it can be performed.

**REFERENCES**


**Unplanned Transport Events and Severity of Illness: Are We Conveying the Whole Picture?**

To the Editor.—

We congratulate Yeager et al on their recently published report, “Pretransport and Posttransport Characteristics and Outcomes of Neonates Who Were Admitted to a Cardiac Intensive Care Unit.”1 A study addressing the outcomes of infants requiring transport to a regionalized pediatric cardiac ICU (CICU) is long overdue. As the American Academy of Pediatrics Section on Cardiology and Cardiac Surgery stated in their guidelines for pediatric cardiovascular centers, pediatric cardiovascular centers must be able provide the “diagnostic services and the full range of treatments, interventions, and surgeries needed to produce high-quality outcomes.”2 We believe that the first step in this continuum of care is the safe and timely transport of out-born neonates to these tertiary care centers. Although transport of these infants is rather common, little is found in the literature regarding the safety and quality of care rendered by transport teams during these transfers.

As we read this report with much interest, several questions arose in terms of the cause of the suboptimal clinical and laboratory values in the out-born group, the authors’ definition of “major transport complications,” and the identification of those patients who were clinically unstable. First, 55 (45%) of the out-born infants had “suboptimal” values on arrival to the receiving CICU. This figure seems high. In addition to being related to the severity of cardiac disease (defined by the Risk Adjustment for Congenital Heart Surgery-1 [RACHS-1] score), can the authors estimate the degree to which these derangements were caused by poor transport care? Although they occurred in both the local and the longer-distance transports, surely not all suboptimal values were inevitable. Were suboptimal values associated with the use of nonpediatric specialty care teams? More importantly, was there an increased mortality rate found in the transported infants who arrived with “out-of-range” clinical parameters when compare to those who did not?
The group also reported no major transport complications in this article. However, they describe 2 patients who were not intubated and had arterial saturation (SpO2) of <50% on arrival to the CICU, problems classified as metabolic deterioration but not unplanned transport events. Both of these patients had normal saturation levels before leaving the outlying facility. We would consider profound hypoxemia to be very much an unplanned transport event, especially in 2 nonintubated patients. If hypoxia of this level is present, should it not be treated, even in infants with congenital heart disease? Clearly, situations such as these are reflective of the team’s performance and not only the patient’s underlying severity of illness.

Finally, we would be very interested in knowing more about the degree of physiologic instability of the cohort. The RACHS-1 score, which categorizes lesion complexity and risk of repair, nicely illustrated that locally born infants suffer from more complex cardiac lesions when compared with out-born infants. However, the RACHS-1 score does not incorporate any physiologic variables, which makes it difficult to assess the degree to which adverse events in the transported patients were related to underlying disease, overall clinical status, or management during transport. For example, did any of the patients have hypotension, evidence of poor cardiac output (eg, decreased urine output, prolonged capillary refill time), or seizures (all of which have been associated with mortality in the newborn population) before or during transport? Reporting the physiologic data would help us to assess and design interventions needed to improve the interfacility transport of critically ill infants with congenital heart disease.

As the pediatric interfacility transport setting becomes a growing area of research interest, we must ensure that we define major transport complications consistently and control as much as possible for severity of illness. By using standard variables and quantifying mortality risk, those who practice in the area of transport medicine will be able to accurately evaluate and meet quality-of-care benchmarks while achieving the best-practice outcomes in children with congenital heart disease.

In Reply.—

We appreciate the comments and insights offered by Kuch et al. Our review was an attempt to characterize the current status of infant cardiac transport, but clearly the data suffer from the retrospective nature of the study. Because these infants were transported by different teams, with differing protocols and inconsistent record-keeping, some salient issues cannot be analyzed.

With regard to the specific points raised, we agree that 45% of transported infants arriving with some suboptimal value is of concern and is an issue that we hope to address going forward. We do not feel that they are “inevitable.” We would consider these data a baseline against which to judge future improvements in transport care.

Similarly, it was not our intent to define “optimal” management or to speculate on whether some additional pretransport intervention would have been appropriate to further stabilize an infant. We were characterizing the transport system and summarizing these infants’ status. The 2 infants described with progressive hypoxemia might have benefited from pretransport intubation, but we would not characterize their course as an in-transport “event.” Perhaps our findings will encourage transport teams and nurseries to reexamine their management protocols. We did look at infants whose selected values fell out of range during the transport. As noted in the article, none of those infants died, and this group was not essentially different as defined by our limited outcome measures.

Vital signs were reported as an aggregate for the groups and did not seem to be predictive of adverse outcome. With regard to other parameters that might have been measured, such as urine output and capillary refill, a prospective study design will be necessary to standardize assessment, documentation, and reporting.

It is our current intention to monitor our transport system in a more rigorous and ongoing fashion and to develop methods for feedback to the transport teams and referring hospitals in an effort to improve the arrival status of these newborns. We would enthusiastically endorse a more refined and standardized set of metrics.

REFERENCES

1. Yeager SB, Horbar JD, Greco KM, Duff J, Thiagarajan RR. Pretransport and posttransport characteristics and outcomes of neonates who were admitted to a cardiac intensive care unit. *Pediatrics*. 2006;118:1070–1077


doi:10.1542/peds.2006-3254

Bradley A. Kuch, RRT-NPS
Department of Critical Care Medicine/Transport
Children’s Hospital of Pittsburgh
Pittsburgh, PA 15213

Ricardo Munoz, MD, FAAP, FCM
Departments of Critical Care Medicine, Pediatrics, and Surgery
Richard A. Orr, MD
Departments of Critical Care Medicine and Pediatrics
R. Scott Watson, MD, MPH
Clinical Research and Systems Modeling of Acute Illness Laboratory
Departments of Critical Care Medicine and Pediatrics
University of Pittsburgh School of Medicine
Pittsburgh, PA 15261

PEDIATRICS Volume 119, Number 3, March 2007 649
for assessing the safety and efficacy of neonatal cardiac transport, and we hope that our findings will stimulate additional discussion as well as regional or national standards of evaluation and outcome.

Scott B. Yeager, MD
Department of Pediatrics
University of Vermont School of Medicine
Burlington, VT 05401

Peter C. Laussen, MBBS
Department of Cardiology
Children’s Hospital Boston
Boston, MA 02115

doi:10.1542/peds.2006-3488

Industry-Sponsored “Expert Committee Recommendations for Acne Management” Promote Expensive Drugs on the Basis of Weak Evidence

To the Editor.—

I have several concerns about the “Expert Committee Recommendations for Acne Management,”¹ published in Pediatrics in September 2006. First, the 2 authors are both associated with various companies that make acne products, including Galderma (which funded the work). Galderma makes Differin (adapalene), which is prominently and favorably discussed in the article. How were the members of the “expert committee” selected? What role, if any, did the sponsor have in their selection?

Second, the recommendations are poorly supported by evidence. The first recommendation listed is: “A topical retinoid should be the foundation of treatment for most patients with acne.” I have always started with benzoyl peroxide, which is available over-the-counter and considerably less expensive. However, not one randomized trial comparing topical retinoids to benzoyl peroxide was cited to justify their recommendation. In fact, the few trials I could find that directly compared the two suggested that, if anything, benzoyl peroxide is superior.²–⁴

Third, the article states that “[s]tudies comparing the potential for cutaneous irritation among topical retinoids have consistently shown that the adapalene molecule is best tolerated of the available retinoids (Fig 7).” Their Fig 7 consists of 2 graphs of “average score” by “measurement points” for 4 topical retinoids and white petrolatum; the legend describes it as: “Tolerability of topical retinoids: 21-day cumulative irritancy study in healthy volunteers.” The average score on this graph is completely uninterpretable. What is the significance of a 1.5-point difference? Why study healthy volunteers rather than acne patients? The legend indicates that data for the figure are on file at Galderma Laboratories. In other words, readers of Pediatrics are being asked to recommend this expensive new medication on the basis of uninterpretable results on inappropriate subjects from an industry-sponsored, unpublished, unreviewed study.

Finally, the expert committee recommendations include no discussion of medication costs. At this writing (October 2006), Drugstore.com is offering 10% benzoyl peroxide gel (ZapZyt) at $4.99 for 30 g, compared with $99 for 45 g of 0.1% generic tretinoin and $121 for 45 g of 0.1% Differin. This sort of information is extremely relevant to practicing pediatricians and their patients with acne. Why was it omitted?

I am accustomed to seeing and discarding drug-promoting, industry-sponsored reviews in throw-away journals or sponsored journal supplements. It was disappointing to see one uncritically published in Pediatrics.

Thomas B. Newman, MD, MPH
Department of Epidemiology
University of California
San Francisco, CA 94143-0560

REFERENCES
3. Belknap BS. Treatment of acne with 5% benzoyl peroxide gel or 0.05% retinoic acid cream. Cutis. 1979;23:856–859

In Reply.—

We acknowledge the concerns that Dr Newman has regarding our recent publication and, in turn, submit this response to each of his criticisms.

The Global Alliance to Improve Outcomes in Acne is made up of 20 dermatologists from around the world. The recommendations that they compiled were the product of in-depth evaluation of the basic science understanding of the pathogenesis of acne and the available clinical trial data, all tempered with their own clinical experience. Almost all major clinical trials for acne were performed by at least 1 member of this group. Most members have performed clinical trials for virtually all of the major manufacturers of prescription acne products, because acne is the primary focus of their careers. None have a single substantial interest in any 1 acne product, retinoid or antimicrobial.

Dr Newman is correct: there are no substantial clinical trials that have compared the efficacy of a retinoid to benzoyl peroxide alone. However, the data supporting
combination therapy, which is what our article advocates, are from well-designed studies that have proven the efficacy of using both benzoyl peroxide and retinoid products together. It should be noted that in adapting the Global Alliance’s recommendations for Pediatrics, the treatment algorithm was modified to include benzoyl peroxide as monotherapy for mild inflammatory acne, because undoubtedly pediatricians see many more cases of early, untreated acne than dermatologists. Recently, a large trial was undertaken by the National Health Service in the United Kingdom to determine efficacy and cost-effectiveness of various topical and oral antimicrobial combinations in the treatment of mild to moderate acne. Although benzoyl peroxide was found to be the most cost-effective and clinically effective treatment, it had the highest irritancy as well. Less irritation was seen when it was combined with topical erythromycin. It is notable, however, that overall improvement was suboptimal; approximately one quarter did not complete the study, and 55% sought additional treatment after the study. Ninety-five percent of the patients still had residual acne at study completion.

Obviously, both efficacy and cost-effectiveness should be considered whenever choosing an acne treatment plan. In practice, we are limited by such factors as tolerability, compliance, insurance coverage, and celebrity endorsements. However, our end goal is the same: to improve outcomes for our patients.

Andrea L. Zaenglein, MD  
Diane M. Thiboutot, MD  
Department of Dermatology  
Pennsylvania State University College of Medicine  
Hershey, PA 17033

REFERENCE


doi:10.1542/peds.2006-3393

Preparing for Asthma-Related Emergencies in Schools

To the Editor.—

Hillemeier et al1 reported in Pediatrics that preparedness for an asthma-related emergency in Pennsylvania schools is suboptimal. Although not cited in the article, similar deficiencies have been identified in New Mexico schools, including the urban and rural school deficiencies. For example, only 20% of New Mexico school nurses reported having nebulizers on campus, and only 45% reported having peak flow meters. As for asthma management plans, 24% of students with asthma did not have a plan, and of those with a plan, only 51% of the nurses either had had input or were identified in the plans.2 In addition, the New Mexico study surveyed for 2 emergency, life-saving medications for severe asthma exacerbations: oxygen and autoinjectable epinephrine. Although 20% of New Mexico schools have oxygen, rural schools were 16 times more likely than urban schools to have it available. Rural schools were also twice as likely to have autoinjectable epinephrine available (life-saving for asthma exacerbations in extremis and food allergies). Both of these medications are well within the capability of school nurses to administer, are of relative low risk in severe situations, and offer additional benefits to β-agonist administration.

Generally, there are 3 major areas of emergency preparedness in schools: emergency response plans (including emergency medical services [EMS] interaction/involvement), equipment, and training. Emergency medical response plans for schools are of 3 distinct types: (1) all hazards (eg, disaster, acts of terrorism); (2) student-specific individual response plans (eg, asthma, seizure); and (3) response plan for general emergencies (eg, playground injury, principal with chest pain). There are components common to each type of plan, but the plans should indicate the unique aspects of response and roles for the school nurse and others. Hillemeier et al reported that 94% of Pennsylvania schools have an emergency response plan, but the extent or type of plan were not reported. Similarly, 86% of New Mexico school nurses reported an emergency plan, but the type and details varied greatly from simply “call 911” to elaborate, detailed plans. In addition, only 44% of school nurses participated in school disaster plans, and only 11% were involved in community disaster planning. Although EMS activation is an integral part of any school emergency plan and involving EMS during preparation efforts is prudent,4 there was great variation in EMS response time and provider training level in rural versus urban schools.1

Finally, training for school nurses and school staff in school emergency preparedness, including asthma management, is paramount. There are a limited number of emergency training opportunities specific to school nurses including the Emergency Medical Services for Children Programs from Connecticut, Illinois, and New Mexico and the National Association of School Nurses. The most recent is a Web-based, 14-hour video-scenario course entitled “Virtual School Nurse and EMS Training.”4 This course contains modules on emergency planning, triage and assessment, respiratory equipment (asthma), and clinical scenarios including asthma.

Unfortunately, the school nurse may not be available on campus when an individual student with an asthma exacerbation presents. It has been reported that school staff may lack knowledge regarding asthma. A program to train nonmedical school staff to recognize signs of respiratory distress in children with asthma has been
described and evaluated. Using video-footage examples of children with asthma in respiratory distress, school teachers were instructed on general asthma signs/symptoms and management, including when to activate EMS. Asthma is but one emergency situation/condition that may arise in the school setting. Preparedness for emergencies in schools is critical, and school nurse and school staff roles in preparedness are crucial.

Robert E. Sapien, MD, FAAP
Department of Emergency Medicine
Division of Pediatric Emergency Medicine
University of New Mexico Health Sciences Center
University of New Mexico
Albuquerque, NM 87131

REFERENCES
1. Hillemeier HM, Gusic ME, Bai Y. Rural and urban children with asthma: are school health services meeting their needs? Pediatrics. 2006;118:1097–1103

doi:10.1542/peds.2006-3335

In Reply.—

We appreciate Dr Sapien’s response to our article on school health services for children with asthma. The limited availability of asthma equipment and the lack of individualized asthma management plans he describes in New Mexico schools are consistent with our findings in Pennsylvania, which suggests that the need for more optimal school asthma management is widespread. Dr Sapien also highlights the importance of overall emergency preparedness in schools, an issue that fortunately is receiving increased attention. It is interesting that in both Pennsylvania and New Mexico, rural schools were found to be more likely to have emergency equipment available than urban schools. This may be a strategic response to perceived barriers including less optimal nurse staffing availability and greater distances to other sources of emergency care. We also agree that training about emergency asthma management is vitally important, for school nurses as well as other school personnel, and are encouraged by recent attention to the development of relevant educational programs, including material specifically designed for rural school settings.

Screening for Developmental Dysplasia of the Hip

To the Editor.—

Representing the Pediatric Orthopaedic Society of North America, we are writing in response to the special article by the US Preventive Services Task Force (USPSTF) “Screening for Developmental Dysplasia of the Hip: Recommendation Statement.” We agree with the USPSTF that the available literature does not offer the highest level of evidence regarding screening and treatment. We also appreciate the difficulties in formulating recommendations for a disease with a severity spectrum that can range from severe teratologic dislocations to mild acetabular dysplasias. However, the authors of this recommendation statement reached their conclusion with flawed evidence synthesis (using “best data” rather than a model-driven method) and a misunderstanding of current practice, particularly as it relates to ultrasound use and the management of mild findings in newborns. Without developmental dysplasia of the hip (DDH) screening, our nation would have thousands more children each year with a preventable disability, a sharp rise.
in the surgical management of hip dysplasia, and a substantially higher liability risk for primary care providers.

In an ideal world, our literature would report several large, prospective, double-blind studies comparing treated and untreated infants, with follow-up for 70 years to document the full disability of a subluxated or dislocated hip. Such studies are both impractical and unethical. The authors of the American Academy of Pediatrics clinical practice guideline were wise enough to recognize this problem, and instead of relying on a data-driven method, used a model-driven method that combined the use of an expert panel, decision modeling, and evidence synthesis. In their flawed model, the USPSTF “was concerned about the potential harms associated with treatment of infants identified by routine screening.” They based their concern on an avascular necrosis (AVN) rate as high as 60% in children who were treated for DDH. The USPSTF failed to recognize that the vast majority of AVN results from treating late-presenting DDH with closed or open reduction. Currently, as a result of infant screening, early detection, and use of the Pavlik harness, the risk of AVN is ≤1% (1% in 546 cases; 0% in 547 cases). A sharply higher risk of AVN would be expected if screening were abandoned and DDH was missed in infants, because surgical open hip reduction would again become commonplace. Those of us caring for children from the underdeveloped world are all too familiar with this reality.

The USPSTF authors also suggest that if the screening process suggests an abnormality, yet none develops, the parents have had to incur unnecessary anxiety and concern over the future of their child. This is a fallacious reason against screening for DDH or any other health care issue in infancy in which early detection can lead to a simple, definitive treatment of a potentially pathologic condition.

We urge the USPSTF to gain a better understanding of current practice and risks before developing future recommendations. Under their “Clinical Considerations,” the USPSTF reported that ultrasonography was one of the most common methods of screening. This is not true in America; instead, periodic physical examination (a harmless test) is our primary hip screening tool. Ultrasound is only used when there are very high risk factors (breech girls) or an abnormal physical examination. Those of us managing DDH are well aware that many early, mild findings spontaneously resolve in the first few months of life. Therefore, we only use the Pavlik harness for newborns with severe dysplasia or infants with persistent dysplasia. For the reasons outlined above, we urge primary care physicians to continue to follow the clinical practice guideline for early detection of DDH outlined by the American Academy of Pediatrics.

In Reply.—

Schoenecker and Flynn fault the US Preventive Services Task Force (USPSTF) in its assessment that there is insufficient information to make a recommendation about routine screening for developmental dysplasia of the hip. They fault the USPSTF for (1) use of a “flawed” methodology that does not include data from models and (2) failure to recognize current clinical practice in terms of the use of surgery (and, therefore, the resultant risk of complications and, thus, the reach of complications thereof) and the choice of ultrasound versus physical examination screening.

The USPSTF is an independent panel charged by Congress with the responsibility of reviewing the scientific evidence for clinical preventive services and develop evidence-based recommendations for the health care community. The USPSTF bases its recommendations on systematic evidence reviews that follow an established and transparent methodology used by all the evidence-based practice centers that conduct these reviews under contract with the Agency for Healthcare Research and Quality (AHRQ). The USPSTF’s methods of summarizing the evidence and deriving a recommendation also adhere to an established and transparent methodology.

REFERENCES

doi:10.1542/peds.2006-2953
Information on the background and methods used by the USPSTF can be seen on the AHRQ Web site (www.ahrq.gov/clinic/uspstfmeth.htm).

In 2006, the USPSTF examined what is known about the natural history of developmental dysplasia of the hip as well as the evidence for accuracy of screening tools, efficacy of treatment, and harms of screening and treatment. The findings of the USPSTF were published in Pediatrics1,2 and are summarized on the AHRQ Web site (www.preventiveservices.ahrq.gov), where the supporting systematic literature review and evidence synthesis can also be found.

Overall, the USPSTF found that the test characteristics of physical examination as a screening tool were unknown because of the absence of data using appropriate criterion standards, and that although screening (whether by physical examination or ultrasound) does lead to earlier detection, most newborn hips identified as abnormal by physical examination or ultrasound resolve spontaneously without therapy. In addition, the USPSTF found that the evidence of effectiveness of surgical and nonsurgical therapy is poor, and that there are significant potential risks of therapy with avascular necrosis rates reported as high as 60% (and, admittedly, as low as 0%). Note that even if the particular risk of avascular necrosis were zero, no surgical procedure can be seen as harmless, given the risks of anesthesia, etc. Using a model to predict outcomes of different screening or treatment scenarios requires stable and accurate estimates regarding the test characteristics of screening tests and the impact of treatment in certain patients. In reviewing the peer-reviewed published literature, the USPSTF was unable to find reliable estimates of these factors. Given the uncertainties throughout the published literature, there was insufficient evidence for the USPSTF to make an evidence-based judgment about net health benefits (benefits minus harms), thus the “I” recommendation.

Although Schoenecker and Flynn point out how many children might go undiagnosed if screening were to stop, and they downplay “unnecessary anxiety and concern” on the part of parents, in fact both the known and potential harms of screening must be balanced against the benefits in making a summary assessment. The USPSTF could not find enough rigor in the data to assess this balance and conclude that there was a net benefit for the service.

Evidence-based methodology is just that; it draws from what has been studied and demonstrated through rigorous scientific study. The USPSTF methodology does not include expert opinion or the anecdote of experience in lieu of evidence gathered in studies in which bias is minimized. When intervening on the entire population of well patients, we should base our recommendations only on the best evidence of benefit available and should not make a recommendation in the absence of such evidence. We appreciate and respect that there are other physicians who feel differently. However, we look at the work of the USPSTF as setting an evidence-based anchor for preventive services, and this is a vital service in informing health policy.

Ned Calonge, MD, MPH
Chair
Diana Petitti, MD, MPH
Vice-chair
US Preventive Services Task Force

REFERENCES

doi:10.1542/peds.2006-3667

Payment for Telephone Care: From Policy to Practice

To the Editor.—

I read with interest the American Academy of Pediatrics (AAP) policy statement titled “Payment for Telephone Care.”1 It is not often that policy statements strike at the core of real problems facing rank-and-file pediatricians. Past AAP Annual Leadership Forum resolutions urged the academy to take a firm stance and deem it appropriate and ethical to charge for telephone care provided by pediatricians. This policy statement made that case eloquently and comprehensively.

Our society’s increasing desire and expectation for access to care at all hours will continue to expand the importance of efficient, high-quality telephone care. Yet, what began as an occasional courtesy on the part of the physician has become a consuming, and almost completely uncompensated, demand on the part of patients and payers.

It is regrettable that although it provided a compelling rationale for payment, the policy statement gave little specific information on how to implement its recommendations.

Thus, it was with some relief that I found the “Payment for Telephone Care” toolkit on the AAP Web site (www.aap.org/moc/index.cfm?view=news [requires AAP member login]). This extensive document provides the specific, practical information necessary for general pediatric offices to operationalize the new policy. Each recommendation is addressed in detail, with helpful suggestions and useful tools such as letter templates and sample documentation forms. I found this online docu-
ment to be one of the most complete and timely aids to policy implementation produced by the academy.

Compensation is appropriate, ethical, and essential to maintain telephone access to pediatricians and assure the quality and sale delivery of telephone care. The recent AAP policy statement outlined the justification, and their toolkit provides the vital bridge from policy to practice. It now falls to us, the practicing pediatricians, to actualize our goals.

Randall S. Sterkel, MD, FAAP
Department of Pediatrics
St Louis Children’s Hospital
St Louis, MO 63141

REFERENCE
doi:10.1542/peds.2006-3267

In Reply.—

As Dr Sterkel correctly points out, the issue of payment for telephone care as addressed in a recently published American Academy of Pediatrics (AAP) policy statement is of great importance to pediatricians. Although pediatric practice entails a great deal of telephone care, physicians in many other medical specialties also experience high demands for telephone care and are concerned that these services, which entail medical decision-making, practice expense, and malpractice risk, are rarely paid.

Over the past several years, there has been increasing support from the AAP and other specialty societies including the American College of Physicians and the American Academy of Family Physicians to revise the existing set of Current Procedural Terminology (CPT) codes under which telephone care services may be reported and paid by private and government payors. The AAP, through the Section on Telephone Care and its Committee on Coding and Nomenclature and with the strong support of the American Academy of Family Physicians and the American College of Physicians, has played a key leadership role in developing a new set of CPT codes for telephone care and other “non–face-to-face” services, including online medical evaluation and management (E/M) services.

These efforts have resulted in significant progress toward revising the existing system of payment for telephone care. At its October 2006 meeting, the American Medical Association’s CPT panel approved 4 new codes for non–face-to-face care, including 3 new codes for telephone care and a new code for online E/M services.

The new telephone codes, which are time based, will replace the existing telephone care codes in the case management section of the CPT manual. These codes may be used to report a telephone E/M service provided by a physician to an established patient, parent, or guardian that did not originate from a related E/M service provided within the previous 7 days or lead to an E/M service or procedure within the next 24 hours or soonest available appointment.

The other new non–face-to-face CPT code may be used to report an online E/M service provided by a physician using the Internet or similar electronic communications network to an established patient, guardian, or health care provider that did not relate to a related E/M service provided within the previous 7 days.

Physicians around the country recently participated in a survey sponsored by the American Medical Association’s Relative Value Scale Update Committee, which will be used to determine the value involved in delivering these non–face-to-face services within the resource-based relative value unit (RB-RVU) system.

The development of new CPT codes for non–face-to-face care represents an important step that supports the delivery of patient care in an increasingly “connected” practice environment. The AAP can be proud of its members for their leadership and advocacy in these efforts.

Sanford M. Melzer, MD, MBA
Medical Administration
Children’s Hospital & Regional Medical Center
Seattle, WA 98105-0371
for the 2005–2006 AAP Section on Telephone Care

REFERENCE
doi:10.1542/peds.2006-3693

Improving Developmental Screening: Combining Parent and Pediatrician Opinions With Standardized Questionnaires

To the Editor.—

We commend Rydz et al for their careful and systematic investigation of the use of parent-completed questionnaires in developmental screening. In the newly revised American Academy of Pediatrics (AAP) policy statement on developmental surveillance and screening, the importance and potential benefit of the use of questionnaires has been recognized in both continuous surveillance and periodic screening at the 9-, 18-, and 24- or 30-month visits. However, conclusions drawn from this study about the utility of these questionnaires must be interpreted carefully in light of methodologic problems and incomplete data.

In this article, data that may shed additional light on
the value of combining the pediatrician’s opinion with parental questionnaires in developmental surveillance and screening were not presented. Although they stated that the pediatrician’s opinion had good specificity but poor sensitivity, no data were shown to support this statement (despite the mention of the data being present in their Table 4). As stated by the authors, it can be concluded that the pediatrician’s opinion may be similar to that of the Child Development Inventory, both of which underidentified affected children. The authors also did not share the potentially useful data obtained from the parents, according to their “Procedures” section, regarding potential developmental concern. The responses to these questions, similar to those incorporated in the Parents’ Evaluation of Developmental Status3 (www.pedtest.com), another well-accepted screening measure, may also be useful when compared with both the pediatrician’s opinion and the questionnaires.

Finally, the questionnaire data and the pediatrician opinion were combined, with cases excluded when the findings were discordant. The new AAP policy statement strongly recommended that a child be referred for additional evaluation when any concern (from a parent, pediatrician, or test result) is identified. By excluding discordant cases, one can expect underidentification of children with developmental problems. In the practice of developmental screening and surveillance, we recommend the incorporation of parent-completed questionnaires or directly administered screening tests into the process of surveillance and screening. However, their results should be combined with attention to parental concerns and the pediatrician’s opinion, rather than replacing them, to augment the screening process and increase identification of children with developmental disorders. As the authors noted, their study highlighted the limits of screening for developmental problems at a single point in time, given the wide variability in developmental pathways, and the potential benefits of serial testing.4 The same could be said about single pieces of clinical information. It is this premise of developmental screening’s limits on which the current algorithm for surveillance and screening was constructed. It is our hope that serial surveillance and screening with data from multiple sources can overcome the limits in sensitivity and specificity of the individual instruments and lead to better identification of the “true” cases of children with developmental disorders.

Although this study looked at a small homogeneous sample of highly educated, middle- to upper-class families, we look forward to additional testing of these techniques in large and diverse samples such as those in the AAP Developmental Surveillance and Screening Policy Implementation Project (www.medicalhomeinfo.org/screening/DPIP.html) and the North Carolina Assuring Better Child Health and Development Academy5 to find optimal techniques for early identification of children with developmental disorders.

Paul H. Lipkin, MD
Chair, AAP Policy Revision Committee for Developmental Surveillance and Screening
Chair, AAP Council on Children With Disabilities
Center for Development and Learning
Kennedy Krieger Institute
Johns Hopkins University School of Medicine
Baltimore, MD 21205

Hilary Gwynn, MD
Neurology and Developmental Medicine
Kennedy Krieger Institute
Johns Hopkins University School of Medicine
Baltimore, MD 21205

REFERENCES
5. Earls MF, Hay SS. Setting the stage for success: implementation of developmental and behavioral screening and surveillance in primary care practice—the North Carolina Assuring Better Child Health and Development (ABCD) Project. Pediatrics. 2006;118(1). Available at: www.pediatrics.org/cgi/content/full/118/1/e183

doi:10.1542/peds.2006-3529

In Reply.—

We appreciate the commendation offered by Lipkin, in his role as Chair of both the American Academy of Pediatrics Council on Children With Disabilities and the Policy Revision Committee for Developmental Surveillance and Screening, and Gwynn. The commendation serves to validate our intent and efforts with respect to our initial attempt to evaluate prospectively 2 parent-completed questionnaires relevant to developmental screening.1 Our efforts were motivated by observations that despite widely endorsed policy statements,2 practice constraints were limiting the widespread use of standardized screening tools,3 with the result being a lamentable late identification of developmentally impaired children.4
To respond to Lipkin and Gwynn’s specific questions, the pediatrician’s opinion was evaluated as a modifier of concerns raised by the Ages and Stages Questionnaire and Child Development Inventory. Specifically, if the pediatrician was concerned by some aspect of the child’s development, did this make it more likely that a failing score on the Ages and Stages Questionnaire or Child Development Inventory would be associated with a failure on the Battelle Development Inventory? The answer was no, thus our statement that the pediatrician’s opinion did not improve the predictive value of the questionnaires tested, which suggested good specificity and poor sensitivity with respect to this opinion. Discordant cases were not excluded from the study, as per Lipkin and Gwynn’s suggestion, but were merely excluded from the analysis of the physician’s opinion as a modifier, because we could not a priori provide a hierarchy of relative value to the questionnaire’s result versus the physician’s opinion. As cited in our article and mentioned by Lipkin and Gwynn, previous studies have shown that parents do consistently provide accurate information regarding their child’s development, hence the rationale for standardized parent-completed questionnaires as a viable option to health care provider–completed questionnaires. Thus, parental opinion was formalized by standardized questioning, and it was this opinion that we analyzed.

We would most certainly concur with Lipkin and Gwynn’s suggestion of the limits of relying on either a single point in time or a single piece of information. What is missing at the moment are data to justify rationally proposing a strategy of when to evaluate and with what mechanisms that meets the challenges of timeliness, cost-effectiveness, and efficiency. What is most likely is that a single strategy may not encompass diverse socioeconomic challenges; thus, future testing that builds on our work and that of others in a variety of practice milieus is clearly necessary.

Michael I. Shevell, MD, CM, FRCP
Division of Pediatric Neurology
Departments of Neurology/Neurosurgery and Pediatrics
Montreal Children’s Hospital-McGill University Health Centre
McGill University
Montreal, Quebec, Canada H3H 6P3

REFERENCES

Support for Addressing the Effects of Public Policy on Same-Gender Couples and Their Children

To the Editor—

The special article “The Effects of Marriage, Civil Union, and Domestic Partnership Laws on the Health and Well-being of Children,”1 published in Pediatrics, is an excellent and important step forward in the publication of unbiased, scientifically sound information on a subject that is often poorly and emotionally misrepresented. The authors of the article presented a thorough, clear, and evidence-based review of the relevant public policy, as well as psychosocial and developmental considerations, and their impact on children. Their conclusions were based on this comprehensive and logical review and reflected the commitment of the American Academy of Pediatrics (AAP) to the “attainment of optimal physical, mental, and social health and well-being for all infants, children, adolescents, and young adults.”2

The importance of this article cannot be ignored. As pediatric practitioners at the intersection of 2 states and the District of Columbia, our patients have different legal rights depending on which state they live in. As a result, we have seen the dramatic impact that legislation can have on the lives of children of same-gender parents. We have seen children negatively impacted and families disrupted because one or both parents were denied legal rights. In contrast, we have seen children continue to thrive in loving homes because the rights of both parents were protected under the law. It is clear to us that our role is to care for and support all children and to work to ensure that every child is able to be cared for by loving parents and has access to a comprehensive medical home. Too often, science takes a backseat to political considerations. This is not acceptable, especially when the health and well-being of children and their families are at stake. In publishing this article, Pediatrics has taken an important step in promoting evidence-based policy.

As pediatricians and active members of the AAP, we are proud to see Pediatrics again promote and reaffirm the health of all children regardless of their background or family circumstances. Every child deserves the same protections and access to high-quality health care. Although a great many health disparities still unfortunately exist, we remain confident in the AAP’s dedication to eliminating these disparities. We applaud the
Thirteen-Year Spontaneous Evolution of Helicobacter pylori Gastritis Acquired During Early Childhood

To the Editor.—

Helicobacter pylori infection is usually contracted in early childhood, before the age of 5 years. The infection causes chronic gastritis that is responsible in adulthood for atrophic gastritis, which could lead to gastric carcinoma. However, there are very few available data about the natural evolution of H pylori gastritis during childhood.

We previously published a prospective 2-year study of 18 asymptomatic H pylori–infected children. The study revealed a slight aggravation of the histologic features of H pylori gastritis despite stable H pylori colonization. No glandular atrophy, intestinal metaplasia, or any other lesions that could lead to gastric carcinoma were found in these children during the 2-year study.

We recently had the opportunity to obtain information on long-term natural evolution of H pylori gastritis in a 15-year-old adolescent girl. She had participated in the prospective study mentioned above when she was 2 years old. Because she was free of symptoms during this 2-year study, no eradication treatment was proposed, and she was then lost to follow-up. She remained free of symptoms during >12 years and came back to us 13 years later because of frequent epigastric pain (twice a week for 3 months), with vomiting but no weight loss. A new endoscopy showed persisting nodular gastritis, and gastric biopsies (n = 5) revealed that she remained infected with H pylori (both at histology and culture); chronic inflammation was intense in antrum and moderate in fundus, with slight activity in both locations. There was neither atrophy nor metaplasia. When comparing sequential gastric biopsies over the 13-year follow-up period, gastritis progressed only slightly: initial histologic features in both locations consisted of moderate inflammation with slight activity. Because her H pylori strain was sensitive to all the antibiotics tested, a 7-day treatment with omeprazole, amoxicillin, and clarithromycin was given, which resulted in eradication of the H pylori 6 weeks later.

This unique case report shows that untreated H pylori infection acquired during early childhood persists without major progression in gastritis. More than the duration of infection, this suggests that exposure to other gastric aggressors and/or cocarcinogens (such as antiinflammatory drugs, special food habits, tobacco, or alcohol) are responsible for gastric atrophy and metaplasia. Moreover, the nonaggravation in histologic features supports recommendations not to screen and treat asymptomatic H pylori–infected children.

Léonie Martigne
Laurent Michaud, MD
Frédéric Gottrand, MD, PhD
Department of Pediatric Gastroenterology
Hepatology and Nutrition
Jeanne de Flandre Hospital
Lille 59037, France
Françoise Boman, MD, PhD
Department of Anatomo-pathology
University Hospital and Medical School
Lille 59037, France
Marie-Odile Husson, MD
Department of Bacteriology
Calmette Hospital
Lille 59037, France

REFERENCES

doi:10.1542/peds.2006-2539

doi:10.1542/peds.2006-3619
ERRATA


An error occurred in one of the immunization schedules accompanying the AAP policy statement “Recommended Immunization Schedules for Children and Adolescents—United States, 2007,” which appeared in the January 2007 issue of Pediatrics (doi:10.1542/peds.2006-3309). In the second immunization schedule, titled “Recommended Immunization Schedule for Ages 7–18 Years,” in footnote 10, the first sentence of the second bullet should read as follows: “Administer 2 doses of varicella vaccine to persons aged <13 years at least 3 months apart.”

doi:10.1542/peds.2007-0232
Infant Care Patterns at Epidemiologic Study of Cystic Fibrosis Sites That Achieve Superior Childhood Lung Function

Raj Padman, MD, Susanna A. McColley, MD, Dave P. Miller, MS, Michael W. Konstan, MD, Wayne J. Morgan, MD, Michael S. Schechter, MD, MPH, Clement L. Ren, MD, Jeffrey S. Wagener, MD, for the Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis

OBJECTIVE. Previous analyses of the Epidemiologic Study of Cystic Fibrosis database revealed that sites with the highest average patient lung function monitor patients and treat with antibiotics more aggressively than those where average lung function is lowest. The aim of this study was to assess whether patterns of care for infants at cystic fibrosis sites with superior average lung function in 6- to 12-year-old children showed any differences from those at the lowest outcome sites.

METHODS. We divided cystic fibrosis sites with ≥20 patients who were 6 to 12 years of age into quartiles on the basis of median forced expiratory volume in 1 second of that age group in 2003 and compared demographic and clinical characteristics and treatment patterns during the first year of enrollment for patients who were aged 0 to 3 years at those sites in 1994 to 1999. The analysis included 755 infants from 12 upper quartile sites and 743 infants from 12 lower quartile sites.

RESULTS. Upper quartile sites had more infants whose disease was diagnosed by family history or newborn screening, fewer infants with symptoms at diagnosis, higher weight for age at enrollment, more white patients, and more ΔF508 homozygotes. Medical conditions and respiratory tract microbiology differed between sites. Infants at upper quartile sites had more office and sick visits; more respiratory tract cultures; and more frequent use of intravenous antibiotics, oral corticosteroids, mast cell stabilizers, and mucolytics; but they received less chest physiotherapy, inhaled bronchodilators, oral nutritional supplements, and pancreatic enzymes.

Key Words
- cystic fibrosis
- early practice patterns
- lung function outcomes

Abbreviations
- CF—cystic fibrosis
- ESCF—Epidemiologic Study of Cystic Fibrosis
- FEV₁—forced expiratory volume in 1 second
- UQ—upper quartile
- LQ—lower quartile

Accepted for publication Sep 18, 2006
Address correspondence to Raj Padman, MD, Division of Pulmonology, Alfred I. duPont Hospital for Children, PO Box 269, Wilmington, DE 19899. E-mail: rpadman@nemours.org

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1414
doi:10.1542/peds.2006-1414
CONCLUSIONS. Both enrollment characteristics and infant care patterns are associated with lung function outcomes in later childhood. Our analysis suggests that pulmonary function of older children may be improved through specific interventions during the first 3 years of life.

The increasing use of newborn screening and prenatal screening has led to more frequent diagnoses of cystic fibrosis (CF) in early infancy. However, there are no published guidelines for care of infants with CF, and little information is available about specific interventions that might improve eventual outcomes. Because nutritional deficits may be observed before 2 months of age, early diagnosis and treatment can prevent malnutrition and improve long-term growth. The nutritional benefits of newborn screening for CF are also well established. Although growth and nutritional indices in early life seem to correlate with lung function later in childhood, the pulmonary benefits of newborn screening are less clear. This may be attributable, in part, to limited knowledge of which interventions during infancy might improve long-term pulmonary outcomes.

The Epidemiologic Study of Cystic Fibrosis (ESCF and ESCF II), a multicenter, longitudinal, observational study, was initiated in 1993 to collect treatment and outcomes data on patients with CF. Wide variation in treatment and outcomes has been observed across sites that participated in the ESCF, and a previous analysis of this database revealed that certain patterns of care were found more often in sites with better pulmonary function outcomes. Because early disease care is likely to have a long-term impact on pulmonary function, we attempted to discern “best practices” for respiratory care by comparing treatment that was prescribed to infants at ESCF sites with the highest median forced expiratory volume in 1 second (FEV₁) in 6- to 12-year-old children with that given at sites with the lowest FEV₁.

METHODS
The ESCF database was used to establish 2 cross-sectional cohorts for this analysis: (1) an outcomes cohort of patients who were 6 to 12 years of age in 2003 and (2) an infant cohort of patients who were treated as infants between 1994 and 1999 at sites that were in the highest or lowest quartiles with respect to FEV₁ in 2003. CF sites that were included in the outcomes cohort reported on at least 20 patients who were aged 6 to 12 years in 2003; 24 sites that participated in ESCF II met these criteria. The best FEV₁ % predicted that was obtained in 2003 was determined for each patient, median FEV₁ was calculated for each site, and those in the highest and lowest quartiles were identified. The infant cohort, aged 0 to 3 years, was used to compare demographic, clinical characteristics, and care patterns identified at upper (UQ) and lower quartile (LQ) sites. All patients who enrolled and participated in ESCF and ESCF II provided informed consent.

Data that were collected for each patient reflected the period from enrollment into the ESCF to 12 months after enrollment. The following demographic characteristics were collected: age, gender, race (non-Hispanic white or other), genotype, and method of CF diagnosis. Clinical characteristics included mean weight-for-age and height-for-age percentiles (during the first year of enrollment), signs and symptoms of lung disease (cough was described as none, occasional, or daily; and the presence of crackles, wheezing, and clubbing was identified using a check box), medical conditions (eg, asthma, sinusitis, elevated liver function results), microbiology and related variables, and nutritional status. The presence of asthma or sinusitis was based on the diagnostic determination of the individual physician. Practice patterns were defined by number of visits and use of parenteral antibiotics and routine therapies.

Continuous variables are presented as means ± SD, and categorical variables are presented as number (%). Two-tailed χ² tests (for categorical variables) and t tests (for continuous variables) were performed to identify any differences between UQ and LQ sites. These variables were examined further by means of multivariable logistic regression modeling, controlling for quartile, genotype, non-Hispanic race, and age. All statistical analyses and summaries were performed using SAS 9.1 (SAS Institute, Cary, NC). P < .05 was considered to be significant.

RESULTS
Quartiles were established from an outcomes cohort of 837 children who were aged 6 to 12 years; 525 of these patients also were included in the infant cohort to evaluate clinical characteristics and practice patterns. A total of 1498 infants were included in the analysis of practice patterns: 755 patients from UQ sites and 743 patients from LQ sites. Figure 1 shows the distribution and overlap of infants and children who were evaluated in the analysis.

Patient Characteristics
Demographic and clinical characteristics for both cohorts are summarized in Table 1. In the outcomes cohort, no differences were observed in mean age or gender distribution between UQ and LQ sites. Median FEV₁ in 2003 was 107.1 ± 3.7% and 89.0 ± 6.5% predicted for UQ and LQ sites, respectively. UQ sites cared for a significantly higher proportion of non-Hispanic white patients (P < .0001) and patients whose CF was diagnosed by newborn screening (P < .0001). UQ sites had fewer infants whose CF was diagnosed by the presence of clinical symptoms (P < .0001). The distribution of other demographic and clinical characteristics was similar be-
were not large. The mean weight-for-age percentiles at the LQ sites (Fig 2), but these differences were not large. The mean weight-for-age percentiles at the UQ and LQ sites were 28.4 and 22.6, respectively, and the mean height-for-age percentiles at the UQ and LQ sites were 31.7 and 25.9, respectively.

Office Visits and Hospitalizations
Infants who received care at UQ sites had more frequent office visits (6.4 ± 3.3 vs 5.9 ± 3.1; P = .009) and sick visits (1.2 ± 1.9 vs 1.0 ± 1.7; P = .014) but a similar number of hospital admissions (0.5 ± 0.9 vs 0.5 ± 1.0) in the 12 months after enrollment (Fig 3).

Treatment
Univariate analyses showed that patients who were cared for at UQ sites received more intravenous antibiotics at home (P < .05), oral corticosteroids (P < .0001), mast cell stabilizer therapies (P < .0001), mucolytics (excluding dornase alfa; P < .0001), and supplemental oxygen (P = .002; Table 5). In contrast, patients from the LQ sites received more airway clearance therapies/chest physiotherapy (P < .0001), inhaled bronchodilators (P = .0001), oral nutritional supplements (P < .05), and pancreatic enzymes (P < .0001). These differences generally were confirmed by multivariable analysis with adjustment for age, race, and genotype. The borderline result for intravenous antibiotics at home (P = .028) no longer was significant after adjustment (P = .059). No differences were noted between groups regarding the use of chronic therapies, including inhaled or oral antibiotics, inhaled corticosteroids, enteral nutrition, or dornase alfa.

Discussion
This analysis compares the demographics, clinical characteristics, and medical care received during the first 12 months after enrollment at ESCF sites in the highest and lowest quartiles for FEV, % predicted in 2003. We found that sites where patients had the best lung function at ages 6 to 12 years had fewer infants with risk factors (nonwhite race, symptomatic diagnosis, P aeruginosa airway infection) compared with sites with lower lung function and that infants who attended these sites had slightly better height and weight percentiles. Some of the risk factors, such as ethnicity and genotype, are not modifiable. Others, including symptomatic diagnosis, cannot be modified at the time of an individual patient’s diagnosis but can be changed in future patient populations with specific interventions. Importantly, specific patterns of patient monitoring and treatment at sites in the highest quartile were significantly different from those at the lower quartile sites.

Comparisons of site outcomes must consider the baseline characteristics of patients who attended those sites, which may not be modifiable by treatment, and differences in treatment approach. Nonmodifiable differences between UQ and LQ sites were seen in this analysis. Fewer ethnic minorities were seen at UQ sites. Minority

Clinical Status
A significantly higher percentage of patients at the LQ sites exhibited cough (P = .001) at ≥75% of clinic visits (Table 2). The incidence of clubbing was similar between groups. The difference in incidence of crackles could not be assessed because of the small number of affected patients.

Medical Conditions
No difference was observed between UQ and LQ sites in the diagnosis of asthma. However, at UQ sites, fewer infants demonstrated abnormal results on liver function tests (P = .003), but more infants received a diagnosis of sinusitis (P < .0001; Table 3).

Microbiology
UQ sites performed airway cultures in a greater proportion of their infants and did more cultures per individual infant than LQ sites (P < .0001 for both). Infants at UQ sites were more likely to have positive cultures for Staphylococcus aureus (P < .0001) and other Gram-negative bacilli (P < .05) and less likely to have positive cultures for Pseudomonas aeruginosa (P < .003; Table 4).

Nutritional Status
Infants at UQ sites exhibited both higher weight-for-age and height-for-age percentiles (P < .0001) compared with patients at the LQ sites (Fig 2), but these differences were not large. The mean weight-for-age percentiles at

![Venn diagram](image-url)
TABLE 1  Demographic Characteristics for Infant and Outcomes Cohort According to Outcomes Quartile (Based on FEV₁ in 2003)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Infant Cohort</th>
<th></th>
<th>Outcomes Cohort</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UQ (N = 755)</td>
<td>LQ (N = 743)</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td>Male 357 (47.7)</td>
<td>387 (52.1)</td>
<td>807</td>
<td>275 (55.3)</td>
</tr>
<tr>
<td></td>
<td>Female 398 (52.7)</td>
<td>357 (47.3)</td>
<td>356 (47.9)</td>
<td>222 (44.7)</td>
</tr>
<tr>
<td>Age, mean ± SD*</td>
<td>11.3 ± 10.80 mo</td>
<td>11.6 ± 10.41 mo</td>
<td>563</td>
<td>7.5 ± 1.51 y</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>Non-Hispanic white 686 (90.9)</td>
<td>599 (80.6)</td>
<td>&lt;.0001</td>
<td>458 (92.2)</td>
</tr>
<tr>
<td></td>
<td>Other 69 (9.1)</td>
<td>144 (19.4)</td>
<td>7.83</td>
<td>39 (7.9)</td>
</tr>
<tr>
<td>FEV₁, % predicted, median</td>
<td>—</td>
<td>—</td>
<td>107.1</td>
<td>75.0 ± 1.47 y</td>
</tr>
<tr>
<td>Genotype, n (%)</td>
<td>ΔF508 homozgyote 359 (47.5)</td>
<td>268 (36.1)</td>
<td>&lt;.0001</td>
<td>210 (42.3)</td>
</tr>
<tr>
<td></td>
<td>ΔF508 heterozygote 250 (33.1)</td>
<td>192 (25.8)</td>
<td>.001</td>
<td>194 (39.0)</td>
</tr>
<tr>
<td></td>
<td>Other 60 (7.9)</td>
<td>56 (7.5)</td>
<td></td>
<td>39 (7.9)</td>
</tr>
<tr>
<td></td>
<td>Unknown 86 (11.4)</td>
<td>227 (30.6)</td>
<td></td>
<td>39 (7.9)</td>
</tr>
<tr>
<td>Diagnosis, n (%)</td>
<td>Clinical symptoms 117 (15.5)</td>
<td>193 (26.0)</td>
<td>&lt;.0001</td>
<td>368 (74.0)</td>
</tr>
<tr>
<td></td>
<td>Family history 142 (18.8)</td>
<td>109 (14.7)</td>
<td>0.032</td>
<td>94 (18.9)</td>
</tr>
<tr>
<td></td>
<td>Screening 33 (4.4)</td>
<td>9 (1.2)</td>
<td></td>
<td>37 (7.2)</td>
</tr>
<tr>
<td></td>
<td>Genotype NA</td>
<td>NA</td>
<td></td>
<td>2 (0.4)</td>
</tr>
<tr>
<td></td>
<td>Unknown NA</td>
<td>NA</td>
<td></td>
<td>5 (1.0)</td>
</tr>
</tbody>
</table>

NA indicates not applicable.

* Age at enrollment for the infant cohort and age in 2003 for the outcomes cohort.

** Diagnosis data were collected from a larger number of categories in ESCF II than ESCF. The term “clinical symptoms” includes acute or persistent respiratory symptoms, edema/hypoproteinemia/hypoalbuminemia, failure to thrive/malnutrition, liver problems, meconium ileus/other intestinal obstructions, nasal polyps/sinus disease, rectal prolapse, and steatorrhea/abnormal stools/malabsorption. Screening includes neonatal screening and prenatal diagnosis.

TABLE 2  Clinical Status for ≥75% of Visits During the First Year After ESCF Enrollment According to Site Outcomes Quartile

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Outcomes Quartile, n (%)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UQ (N = 755)</td>
<td>LQ (N = 743)</td>
</tr>
<tr>
<td>Cough</td>
<td>342 (45.3)</td>
<td>402 (54.1)</td>
</tr>
<tr>
<td>Crackles</td>
<td>3 (0.4)</td>
<td>7 (0.9)</td>
</tr>
<tr>
<td>Clubbing</td>
<td>32 (4.2)</td>
<td>25 (3.4)</td>
</tr>
</tbody>
</table>

Site outcomes quartiles were determined on the basis of ranking the sites according to their median FEV₁, % predicted in 2003 among 6- to 12-year-old patients who were enrolled in ESCF.

TABLE 3  Medical Conditions During the First Year After ESCF Enrollment According to Site Outcomes Quartile

<table>
<thead>
<tr>
<th>Medical Condition*</th>
<th>Outcome Quartile, n (%)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UQ (N = 755)</td>
<td>LQ (N = 743)</td>
</tr>
<tr>
<td>Asthma</td>
<td>160 (21.2)</td>
<td>139 (18.7)</td>
</tr>
<tr>
<td>Elevated liver function tests</td>
<td>19 (2.5)</td>
<td>41 (5.5)</td>
</tr>
<tr>
<td>Sinusitis</td>
<td>146 (19.3)</td>
<td>71 (9.6)</td>
</tr>
</tbody>
</table>

Site outcomes quartiles were determined on the basis of ranking the sites according to their median FEV₁, % predicted in 2003 among 6- to 12-year-old patients who were enrolled in ESCF.

* Data in this table reflect any of these conditions ever reported.

ethnicity is associated with a lower socioeconomic status in the United States, which, in turn, is associated with poorer pulmonary function in patients with CF. Because socioeconomic data were not available for this analysis, it is not possible to evaluate the impact of socioeconomic status separately. Fewer pancreatic-insufficient patients, as identified by pancreatic enzyme use, were seen at UQ sites. Pancreatic insufficiency is associated with increased disease severity.¹⁰,¹¹

Some of the risk factors that were seen at LQ sites can be modified in future populations of infants with CF. Infants from LQ sites demonstrated a greater severity of illness and a lower weight for age at enrollment. Both
weight for age and severity of illness during the first year after diagnosis affect morbidity and mortality in children with CF. These risk factors can be modified by widespread implementation of newborn screening programs. Observational and randomized studies have demonstrated that early diagnosis and treatment improve long-term nutritional status and lung health and are the basis for the Centers for Disease Control and Prevention and the Cystic Fibrosis Foundation consensus for routine newborn screening for CF. In the United Kingdom, infants whose CF is diagnosed by newborn screening require less intensive therapies than those whose CF is diagnosed clinically. Our data show that ESCF sites whose 6- to 12-year-old patients have the highest pulmonary function had more infants whose CF was diagnosed by newborn screening or family history than sites with the lowest pulmonary function. These infants also seem to require less intensive therapy, because less frequent prescription of airway clearance therapy, inhaled bronchodilators, and enteral supplemental feedings were seen at UQ sites. There are no randomized, controlled studies of these therapies in CF, and the association between use of these therapies in infants and outcomes cannot be assessed adequately with the current analysis.

More patients from the LQ sites had sputum cultures positive for *P. aeruginosa*, which is associated with greater airway inflammation and poorer lung function in ensuing years. In contrast, UQ sites performed more cultures per patient, which might be expected to allow earlier detection and treatment. Infection control practices and early intervention for *P. aeruginosa* infection could reduce the frequency of positive *P. aeruginosa* cultures in the population and further improve outcomes.

For achievement of the best health outcomes for infants whose CF is diagnosed without clinical symptoms and for improvement of outcomes for those whose CF is diagnosed with symptoms, it is essential to recognize and apply monitoring and treatment strategies that may improve subsequent health status. Although significant information indicates that important physiologic aberrations occur early in CF, very few data give the clinician insight into the best treatment strategies. For example, there is evidence for early airway infection and inflammation in infants and young children with CF, and structural airway abnormalities are apparent on high-resolution computed tomography scan in infants and young children with CF. Abnormalities that are seen on high-resolution computed tomography are prevalent even in a cohort of children with normal lung function, the conventional marker of lung health. Furthermore, early malnutrition has deleterious effects on cognitive function in children with CF; minimizing the duration of vitamin E deficiency may be associated with better cognitive function. Therefore, early diagnosis and prompt care, including close attention to growth and vitamin status, of infants with CF are expected to improve long-term outcomes. We do not know, however, which interventions might prevent bronchiectasis or slow its progression or whether specific strategies for nutrition in infants and young children could improve further later nutrition.

Although our analysis showed that sites with the highest FEV₁ had a lower proportion of high-risk infants, it is important and most useful to examine the differences in treatment approaches that were used during infancy. In older children, significant differences in care patterns and outcomes have been noted at ESCF sites. More visits to the site, more frequent respiratory tract cultures, and more frequent use of intravenous antibiotics have been documented at UQ versus LQ sites. The present analysis shows that these specific care patterns also are seen in the care of the youngest patients at sites that later demonstrate the best pulmonary function outcomes. Despite lesser severity of illness, infants in the UQ sites were evaluated more frequently (office and sick visits), cultured more frequently, and treated more often with intravenous antibiotics (at home), oral corticosteroids, mast cell stabilizers, and mucolytics than those at LQ sites. Increased use of anti-inflammatory therapies in young children in UQ sites suggests consideration of evaluating the benefits of these agents in prospective, controlled clinical trials. Also surprising was the finding...
that patients at UQ sites received supplemental oxygen more frequently than those at LQ sites; given that UQ patients had a lower risk for severe lung disease, this finding might represent a variability in prescribing practices. The more frequent diagnosis of sinusitis at the UQ sites (P < .001) may be a marker for more frequent antibiotic therapy, leading to better lung function in subsequent years. The significance of this finding is difficult to determine because case report forms captured physician-diagnosed sinusitis but did not give specific criteria for this diagnosis.

It is intriguing that airway clearance and the administration of bronchodilators, interventions that the majority of CF care providers initiate early and consider to be effective, were used more frequently at the LQ than the UQ sites. It is possible that providers at UQ sites do not prescribe these therapies in asymptomatic or minimally symptomatic infants. It also is possible that UQ sites more frequently diagnose CF by newborn screening and that these patients receive less intense treatment.16 Alternatively, this finding may be explained by the variability in defining actual versus prescribed care. The case report forms were designed to capture prescribed therapy, but because case reports are extracted from the medical chart, it is possible that intended treatment, rather than actual treatment, has been captured and that parents of asymptomatic or minimally symptomatic infants are less likely to adhere to airway clearance regimens. This issue might be investigated further through a detailed survey of CF care provider practices.

It is impossible to determine fully the influence of care in previous years versus more recent or current care on pulmonary function outcomes. However, given the early onset of inflammation and bronchiectasis in CF and the progressive nature of CF lung disease, it is compelling to consider the benefits of frequent monitoring and more frequent use of antibiotics in infants and young children as strategies for improving long-term health outcomes. Our study design is limited by the use of 2 cross-sectional cohorts at identical sites rather than a true longitudinal assessment of individual patients. However, 64% of the 6- to 12-year-old patients who were used to define the outcomes cohorts were part of the infant cohort at the same care site. Nonetheless, we believe that our findings are valid. A longitudinal cohort study would be confirmatory, and multivariable analysis then would allow the ascertainment of the relative contribution of patient risk factors and treatment patterns in relation to pulmonary outcomes.

It is important to realize that an epidemiologic study can identify only associations and, as such, does not define cause and effect but rather points out opportunities for future evaluation. In the present study, observation of better lung function at ESCF sites in patients who were aged 6 to 12 years is associated with specific patient characteristics and clinical care patterns in infants in previous years. Although some patient characteristics are not modifiable, others can be modified through currently available strategies such as widespread implementation of newborn screening. Higher FEV₁ is associated with specific practice patterns, including more frequent visits, more sputum cultures, and more frequent anti-

### Table 5: Treatments During the First Year After ESCF Enrollment According to Site Outcomes Quartile

<table>
<thead>
<tr>
<th>Treatment*</th>
<th>Outcomes Quartile, n (%)</th>
<th>UQ (N = 755)</th>
<th>LQ (N = 743)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intravenous antibiotics (at home)</td>
<td>34 (4.5)</td>
<td>18 (2.4)</td>
<td>.028</td>
<td>.059</td>
</tr>
<tr>
<td>Intravenous antibiotics (at home or in hospital)</td>
<td>183 (24.2)</td>
<td>194 (26.1)</td>
<td>.40</td>
<td>.50</td>
</tr>
<tr>
<td>Inhaled antibiotics</td>
<td>72 (9.5)</td>
<td>82 (11.0)</td>
<td>.34</td>
<td>.12</td>
</tr>
<tr>
<td>Oral antibioticsb</td>
<td>581 (77.0)</td>
<td>553 (74.4)</td>
<td>.25</td>
<td>.66</td>
</tr>
<tr>
<td>Airway clearance techniques/ chest physiotherapy</td>
<td>655 (86.8)</td>
<td>697 (93.8)</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Domase alfa</td>
<td>62 (8.2)</td>
<td>77 (10.4)</td>
<td>.15</td>
<td>.072</td>
</tr>
<tr>
<td>Bronchodilator, oral</td>
<td>61 (8.1)</td>
<td>53 (7.1)</td>
<td>.49</td>
<td>.39</td>
</tr>
<tr>
<td>Bronchodilator, inhaled</td>
<td>577 (76.4)</td>
<td>626 (84.3)</td>
<td>.0001</td>
<td>.0005</td>
</tr>
<tr>
<td>Corticosteroid, oral</td>
<td>215 (28.5)</td>
<td>130 (17.5)</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Corticosteroid, inhaled</td>
<td>141 (18.7)</td>
<td>115 (15.5)</td>
<td>.10</td>
<td>.28</td>
</tr>
<tr>
<td>Mast cell stabilizer</td>
<td>260 (34.4)</td>
<td>143 (19.2)</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Nutrition, oral supplements</td>
<td>273 (36.2)</td>
<td>313 (42.1)</td>
<td>.018</td>
<td>.011</td>
</tr>
<tr>
<td>Nutrition, enteral</td>
<td>72 (9.5)</td>
<td>77 (10.4)</td>
<td>.59</td>
<td>.70</td>
</tr>
<tr>
<td>Pancreatic enzymes</td>
<td>720 (95.4)</td>
<td>734 (98.8)</td>
<td>&lt;.0001</td>
<td>.0003</td>
</tr>
<tr>
<td>Mucolytics</td>
<td>71 (9.4)</td>
<td>8 (1.1)</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Oxygen</td>
<td>48 (6.4)</td>
<td>22 (3.0)</td>
<td>.0018</td>
<td>.0008</td>
</tr>
</tbody>
</table>

* Site outcomes quartiles were determined on the basis of ranking the sites according to their median FEV₁ % predicted in 2003 among 6- to 12-year-old patients who were enrolled in ESCF.

b Data that are n (%) reflect any reported usage of these treatments.

P<sub>c</sub> value in first column obtained from 2 test. P<sub>c</sub> value in second column obtained from multivariable logistic regression model controlling for quartile, non-Hispanic white, genotype, and age in months.

Does not include oral quinolones.
otic therapy. The correlations between these practices and improved outcomes generate hypotheses for prospective research and suggest specific strategies that can be implemented in quality improvement initiatives. Frequent monitoring and use of antibiotics may preserve long-term lung health in CF.

ACKNOWLEDGMENTS
This study was supported by Genentech, Inc.

This study was conducted with the ESCF database.

We acknowledge the contributions of the North American Scientific Advisory Group; the investigators and study coordinators of the ESCF; and Jill Luer, PharmD, for editorial assistance.

REFERENCES


Inaccuracy in Pediatric Outpatient Blood Pressure Measurement

Amber Podoll, MD, Michelle Grenier, MD, Beth Croix, RN, Daniel I. Feig, MD, PhD

Cardiology and Renal Sections, Department of Pediatrics, Baylor College of Medicine, Houston, Texas; Renal Section, Department of Nursing, Texas Children’s Hospital, Houston, Texas

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

ABSTRACT

OBJECTIVE. Hypertension is common in the pediatric population. There is increasing evidence for early hypertensive target organ damage that may lead to substantial long-term morbidity. Because a critical aspect of any screening program for hypertension is the ability to measure blood pressure accurately, we compared typical blood pressure measurements at a vital sign station with those that were obtained following recommendations set forth in “The Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents.”

METHODS. We compared the blood pressure measurements that were obtained with standard practice vital sign station screening with those that were obtained by trained personnel in accordance with Fourth Task Force recommendations. A total of 390 children were evaluated at 580 visits to the Pediatric Hypertension Clinic at Texas Children’s Hospital.

RESULTS. Seventy-four percent of the readings were higher at the vital sign station, and only 12% differed by <5 mm Hg for both systolic blood pressure and diastolic blood pressure. The mean difference between vital sign station and examination room was 13.2 ± 8.9 mm Hg for systolic blood pressure and 9.6 ± 7.6 mm Hg for diastolic blood pressure. Multiple regression analyses revealed that age, gender, race, obesity, first versus subsequent visit, essential versus secondary, or white coat hypertension and antihypertensive medications made no statistically significant difference in the lack of correlation of the readings.

CONCLUSION. These results suggest that if pediatricians use vital sign station screening for blood pressure, children with elevated initial measurements must be reevaluated in the examination room.
Hypertension affects >30% of the world population and is a primary risk factor for cardiovascular disease, stroke, and premature death. Numerous studies demonstrate that the target organ damage can be ameliorated and the risk for future disease can be decreased dramatically by early and effective blood pressure (BP) control. Implicit in any program to identify and evaluate patients for hypertension is the ability to measure BP accurately.

Although less common than in adults, between 2% and 4% of children have hypertension. Overt cardiovascular disease and end-stage renal disease secondary to hypertension are very rare in children, but evidence for target organ damage is not. Studies of newly hypertensive children demonstrate that as many as 30% have increased left ventricular mass and >5% have overt left ventricular hypertrophy. Fifteen percent to 20% of hypertensive children have proteinuria, and a recent report suggested that hypertensive children have a variety of reversible somatic symptoms. Autopsy data suggest that children with traditional cardiovascular risk factors have accelerated atherosclerosis, an ominous indication for future disease.

Because of the potential for severe long-term morbidity and the rising prevalence of hypertension in the young, it is becoming increasingly important to screen and evaluate children who are at risk for hypertension. The fundamental component of any screening effort is the act of measuring BP and interpreting its significance. We compared, in a large outpatient population, the standard practices for pediatric BP screening with BP assessment in strict accordance with the best practice guidelines put forth by the Fourth Task Force on the Diagnosis and Management of High Blood Pressure in Children and Adolescents.

METHODS

Patients
We retrospectively analyzed the charts of all children who were referred to the Pediatric Hypertension Clinic at Texas Children’s Hospital between January 2003 and December 2004. A total of 390 children were seen in a total of 580 visits.

Clinical Setting and Staff
Medical assistants were employees of Texas Children’s Hospital and received standard institution training and orientation in clinic care, family relations, and vital sign measurement. This included instruction in BP cuff size selection and appropriate positioning of the patients. The 4 clinic medical assistants each had >2 years of experience at the beginning of the study period, and there was no personnel turnover during the evaluation period. Being a retrospective study, it was not possible to validate the technique of individual technicians before or during the study period.

Vital Sign Station BP Measurement
On children’s arrival at the clinic, a clinic medical assistant assessed children in a vital sign station (VSS) that was adjacent to the waiting room. The clinic routine was to measure height and weight immediately followed by simultaneous measurement of BP and temperature in a chair next to the scale. BP was measured with a Dinamap (GE Healthcare, Fairfield, CT) oscillometric device with the patient seated. The clinic used 6 different Dinamap devices. In January 2003, we used 2 Dinamap 6100 and 2 Dinamap 8100 monitors. During the study period, the 2 Dinamap 6100 monitors routinely were replaced with new Dinamap Pro 400VS monitors because they had been in service for 5 years. The monitors are tested and maintained by the Biomedical Instruments Division at Texas Children’s Hospital. Routine maintenance includes testing and calibration every 3 months. Reported measurements were the mean of 2 upper arm BPs. The medical assistants were instructed in selection of cuff size at the time of their orientation, but there was no independent assessment of the accuracy of their selections. When the pressures differed by >10 mm Hg, a third BP was measured and the reported pressure was the mean of the closer 2 values. BPs of obviously fearful or crying children were not included for the analysis.

Examination Room BP Measurement
Children were seated in a quiet examination room (EXR) for at least 10 minutes before measurement of their BP. All BPs were measured by trained and certified staff and done in accordance with Fourth Task Force recommendations of technique and cuff sizes. BP was measured with Welch Allen (Skaneateles Falls, NY) aneroid sphygmomanometers that were calibrated with mercury devices every 2 months. BPs were measured 4 times at initial visits and twice at subsequent visits. When pressures differed by >10 mm Hg, an additional BP was measured and the reported value was the mean of the closer values. BPs of obviously fearful or crying children were not included for the analysis.

Definition of hypertension is in accordance with the “Fourth Report on the Diagnosis, Evaluation, and Treatment of High Blood Pressure in Children and Adolescents.” Evaluation and diagnosis of children with hypertension in our program has been described previously. Children with initial BP in the pre- to stage 1 hypertension range had history; physical examination; urinalysis; renal ultrasound; ambulatory BP monitoring (ABPM; with 90217 monitor [Space Labs, Redmond, WA]); and laboratory evaluation that included complete blood count, electrolytes, serum urea nitrogen, creatinine, glucose, uric acid, thyroid function tests, fasting
lipids, aldosterone, and direct renin. Echocardiogram was performed in all children with confirmed hypertension. Children with stage 2 hypertension at initial evaluation did not have ABPM, to avoid delay of therapy, and in addition to the studies that were done for stage 1 had renal perfusion scanning for renal scars; noninvasive renal angiography (generally computed tomography); and 24-hour urine collection to screen for pheochromocytoma, Liddle syndrome, apparent mineralocorticoid excess, and glucocorticoid remediable aldosteronism. Children received a diagnosis of essential hypertension when they had confirmed hypertension by Task Force and ABPM criteria for stage 1 hypertension and Task Force criteria alone for stage 2 hypertension and the laboratory and imaging tests all were normal. Children received a diagnosis of secondary hypertension when their evaluation yielded a specific cause for their hypertension. Children who had apparent stage 1 hypertension in clinic and had normal BP by ABPM criteria received a diagnosis of white coat hypertension (WCH).

A small number of children who were referred to the Hypertension Clinic had BPs that fell into range of prehypertension at initial evaluation. These children had ABPM and were grouped with those who had WCH, if normal BP by ABPM criteria, because all children who were referred to the Hypertension Clinic had BPs that fell into range of prehypertension.

Statistical Analysis
Primary end points were the difference between VSS systolic BP (SBP) and EXR SBP and the difference between VSS diastolic BP (DBP) and EXR DBP. Secondary end points were the BP differences in the subgroup populations separated by gender, race, age (in 6-year groups), BMI (normal weight, overweight, and obese), and first or subsequent visit. Statistical significance testing of the continuous variables was by Student’s t test. Analysis of variance was used to assess the differences between subgroups and the whole population. Correlation between VSS and EXR BP measurements was assessed by Pearson calculations. Statistical analyses were performed using Statistica 7.0 software (StatSoft Inc, Tulsa, OK).

RESULTS
Population Characteristics
The characteristics of the study population are shown in Table 1. The population was a pediatric cohort with an age range from 1 month to 18 years (mean: 12.9 ± 4.2 years). It was similar ethnically to the Houston metropolitan area with the exception of fewer children of Asian descent (35% black, 40% white, 23% Hispanic). There were more boys than girls (228 and 162, respectively), and the diagnoses of essential hypertension, secondary hypertension, and the combination of prehypertension and WCH were represented evenly. Slightly more than half of the population, 53.9%, was overweight (BMI >85th percentile for age), and 25.6% were obese (BMI >95th percentile).

### BP Measurements
The BP data for the population is shown in Tables 2 and 3. In Table 2 the SBPs and DBPs shown are the means for all patient BPs at each site, VSS or EXR, whereas the ΔSBP and ΔDBP are calculated by subtracting the EXR BP from the VSS BP for each visit and determining the mean of the differences. Because some EXR BPs were greater than VSS BPs and there was wide variability in the difference of the age and the size of the patients in the population, the ΔSBP does not equal the difference of the mean VSS SBP and the mean EXR BP. At the time of the initial visit, 23 (5.9%) of 390 children had VSS SBPs and DBPs <95th percentile, 6 (26% of those normotensive at VSS) of whom had SBP or DBP >95th percentile in the EXR. Sixty-six children (17% of the new patient visits, including 17 children who had normal VSS BPs) had EXR SBPs and DBPs <95th percentile. A total of 55 (14.1%) of 390 of the children were misclassified as hypertensive or normotensive and another 81 were misdiagnosed as to stage of hypertension by the VSS techniques.

<table>
<thead>
<tr>
<th>Table 1: Study Population (N = 390)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Age, mean, y</td>
</tr>
<tr>
<td>Race, %</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>White</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Gender, %</td>
</tr>
<tr>
<td>Boys</td>
</tr>
<tr>
<td>Diagnosis, %</td>
</tr>
<tr>
<td>Essential hypertension</td>
</tr>
<tr>
<td>Secondary hypertension</td>
</tr>
<tr>
<td>WCH</td>
</tr>
<tr>
<td>Weight, %</td>
</tr>
<tr>
<td>Overweight (BMI &gt;85th percentile)</td>
</tr>
<tr>
<td>Obese (BMI &gt;95th percentile)</td>
</tr>
<tr>
<td>SBP Patient Visits, %</td>
</tr>
<tr>
<td>Child on medication</td>
</tr>
</tbody>
</table>

The mean differences for BPs at VSS and EXR were 13.2 ± 8.9 and 9.6 ± 7.2 mm Hg for SBPs and DBPs, respectively (P > .0001 for both). Subgroups of the population separated by gender, race, age (by 6-year intervals), type of hypertension, body habitus, or first versus subsequent visits revealed ΔSBP ranging from 11.9 to 14.2 mm Hg and ΔDBP ranging from 8.8 to 11.8...
mm Hg, none of which was statistically significantly different for the ΔSBP and ΔDBP for the whole population.

Because of the observation that when serial BPs are measured in children the first BP often is somewhat higher, we repeated the analysis, excluding the first VSS measurement. Table 3 compares the last VSS measurement with the mean EXR data. The mean difference in SBP and DBP are 12.9 ± 10.9 and 9.5 ± 7.1 mm Hg, respectively, differences that are not statistically different from those using mean VSS measurements. When the subgroups gender, race, age, diagnosis, BMI range, and visit number were compared, the only statistically significant differences between the values in Tables 3 and 2 are that the ΔSBP for white patients and the ΔDBP for black patients were smaller when only the last VSS measurement was used. Both values still were statistically significantly different from EXR measurements (P < .001 for both). The similarity in VSS values likely is

<table>
<thead>
<tr>
<th>Parameter</th>
<th>n</th>
<th>Second VSS SBP</th>
<th>Second VSS DBP</th>
<th>EXR SBP</th>
<th>EXR DBP</th>
<th>ΔSBP</th>
<th>P</th>
<th>ΔDBP</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>All visits</td>
<td>580</td>
<td>136 ± 16</td>
<td>78 ± 13</td>
<td>128 ± 16</td>
<td>75 ± 11</td>
<td>12.9 ± 10.1</td>
<td>&lt;.0001</td>
<td>9.5 ± 8.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>341</td>
<td>139 ± 16</td>
<td>77 ± 11</td>
<td>129 ± 16</td>
<td>73 ± 10</td>
<td>12.9 ± 10.9</td>
<td>&lt;.0001</td>
<td>9.5 ± 7.1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Girls</td>
<td>239</td>
<td>136 ± 16</td>
<td>79 ± 14</td>
<td>126 ± 17</td>
<td>76 ± 13</td>
<td>12.5 ± 8.9</td>
<td>&lt;.001</td>
<td>9.6 ± 7.4</td>
<td>.001</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>202</td>
<td>140 ± 13</td>
<td>84 ± 13</td>
<td>129 ± 16</td>
<td>75 ± 12</td>
<td>10.9 ± 7.7</td>
<td>&lt;.0001</td>
<td>8.2 ± 6.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>White</td>
<td>234</td>
<td>138 ± 16</td>
<td>78 ± 12</td>
<td>127 ± 16</td>
<td>73 ± 11</td>
<td>11.7 ± 8.1</td>
<td>&lt;.0001</td>
<td>9.9 ± 7.0</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hispanic</td>
<td>133</td>
<td>137 ± 18</td>
<td>78 ± 14</td>
<td>128 ± 16</td>
<td>75 ± 12</td>
<td>12.6 ± 9.5</td>
<td>&lt;.0001</td>
<td>9.9 ± 8.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–6</td>
<td>58</td>
<td>120 ± 19</td>
<td>73 ± 13</td>
<td>107 ± 19</td>
<td>64 ± 13</td>
<td>14.9 ± 8.9</td>
<td>&lt;.01</td>
<td>11.5 ± 8.2</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>7–12</td>
<td>126</td>
<td>134 ± 16</td>
<td>76 ± 12</td>
<td>124 ± 14</td>
<td>73 ± 11</td>
<td>12.4 ± 9.7</td>
<td>&lt;.01</td>
<td>8.8 ± 7.0</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>12–18</td>
<td>396</td>
<td>142 ± 15</td>
<td>79 ± 14</td>
<td>132 ± 14</td>
<td>76 ± 11</td>
<td>12.7 ± 10.5</td>
<td>&lt;.0001</td>
<td>9.7 ± 7.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WCH/prehypertension</td>
<td>127</td>
<td>127 ± 15</td>
<td>71 ± 10</td>
<td>114 ± 13</td>
<td>67 ± 9</td>
<td>13.5 ± 9.4</td>
<td>&lt;.01</td>
<td>10.3 ± 7.6</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Essential</td>
<td>238</td>
<td>142 ± 14</td>
<td>79 ± 12</td>
<td>132 ± 14</td>
<td>76 ± 10</td>
<td>12.2 ± 10.7</td>
<td>&lt;.001</td>
<td>9.5 ± 7.0</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Secondary</td>
<td>215</td>
<td>137 ± 19</td>
<td>78 ± 14</td>
<td>126 ± 18</td>
<td>74 ± 13</td>
<td>13.5 ± 9.6</td>
<td>&lt;.001</td>
<td>9.3 ± 7.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>214</td>
<td>136 ± 20</td>
<td>79 ± 15</td>
<td>125 ± 19</td>
<td>74 ± 13</td>
<td>14.0 ± 12.8</td>
<td>&lt;.0001</td>
<td>10.6 ± 8.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Overweight</td>
<td>366</td>
<td>139 ± 16</td>
<td>75 ± 10</td>
<td>130 ± 14</td>
<td>74 ± 10</td>
<td>13.1 ± 8.1</td>
<td>&lt;.0001</td>
<td>7.5 ± 5.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Obese</td>
<td>257</td>
<td>141 ± 15</td>
<td>78 ± 11</td>
<td>131 ± 14</td>
<td>75 ± 10</td>
<td>12.4 ± 8.5</td>
<td>&lt;.0001</td>
<td>9.2 ± 6.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Visit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>390</td>
<td>137 ± 18</td>
<td>77 ± 13</td>
<td>131 ± 18</td>
<td>76 ± 12</td>
<td>12.4 ± 11.3</td>
<td>&lt;.0001</td>
<td>9.4 ± 7.4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Later</td>
<td>190</td>
<td>140 ± 16</td>
<td>80 ± 12</td>
<td>124 ± 13</td>
<td>72 ± 9.8</td>
<td>13.5 ± 8.8</td>
<td>&lt;.0001</td>
<td>9.7 ± 7.1</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
because when the 2 VSS measurements were >10 mm Hg different, a third measurement was performed and the closer 2 were averaged (see “Methods”). This practice eliminated many of the high outlier values that often are seen as initial BP measurements.

Analysis of variance for variables that affected ΔSBP revealed statistically significant effects (P < .05) only for VSS SBP and VSS DBP and no significant effect of age, gender, race, diagnosis, laboratory parameters (uric acid, hemoglobin, mean cell volume, creatinine, and potassium), body habitus (height, weight, body surface area, BMI, and BMI percentile), glomerular filtration rate, birth weight, or prescription of antihypertensive medications. Analysis of variance for variables that affected ΔDBP revealed statistically significant effects (P < .05) only for VSS SBP, VSS DBP, and serum potassium (P = .043). The magnitude ΔSBP and ΔDBP was not statistically significantly different for essential or secondary hypertension or WCH.

Correlation of BP Measurements

Figure 1 shows the correlation between VSS and EXR BPs. Figure 1A shows the linear relationship between VSS and EXR SBP with a calculated correlation coefficient of \( r = 0.7199 \). There seems to be somewhat less disparity at the highest BPs, but this was not statistically significant for more than or <160 mm Hg. Figure 1B shows the relationship between VSS and EXR DBPs. The correlation coefficient is only \( r = 0.5947 \), reflecting a wider disparity between measurements than seen in SBP.

DISCUSSION

The goal of the study was to compare BPs that were obtained in a typical pediatric practice setting with the more rigorous methods that are used for the establishment of pediatric normative BP data and for pediatric hypertension trials. Because hypertension in children is defined by comparison with existing normative tables, dramatic differences in technique may confuse the diagnosis. The analysis that we performed does not compare 1 device with another or 1 operator with another but rather the BP measurements that are obtained by standard practice of nonspecifically trained medical assistants using an oscillometric device at a central VSS with those that are obtained by rigorous adherence to Task Force recommendations. The latter practice is comparable to the measurements that are used to generate the pediatric normative data that are used for the diagnosis of hypertension in children.

The magnitude of the differences in VSS and EXR measurements is large. With a mean difference >13 mm Hg for SBP and 9 mm Hg for DBP, the technique differences can lead to substantial misdiagnosis: normotensive children being identified as hypertensive or children with stage 1 hypertensive being identified as having stage 2. In most children, the VSS practices led to an overestimation; however, application of a correction factor, such as subtraction of 13 from SBP, would be unwise because 24% of VSS BPs actually were lower than EXR values. Furthermore, the variance of the differences between measurement techniques was large, rendering correction factors useless.

The degree of difference between VSS and EXR measurements was preserved through all analyzed subgroups. Gender, race, age, and degree of obesity had no significant effect on the differences in SBP or DBP. The lack of impact of age or BMI suggests that incorrect cuff size use is unlikely to explain sufficiently systematic overestimation of BP. We also did not see evidence for an exaggerated WCH effect at the VSS because children with WCH had similar differences in SBP and DBP to all other groups.

This study has several important limitations. First, it was performed at a single center. That it was performed in a hypertension clinic, where there is an emphasis on the measurement of BPs, might be expected to increase the consistency of VSS measurements compared with clinics in which BP is measured less regularly. Second,
we did not control for cuff size selection. It theoretically is possible that systematic use of smaller cuffs at the VSS contributes to the differences seen; however, this requires that the 4 medical assistants were consistent in such errors during the 2-year study period. If this were to be the case, then it further emphasizes the importance of training and technique in the ascertainment of pediatric BP data. The third limitation is that the VSS BPs were always done first, rather than having randomized order to the BP techniques. A decline in BP values with repeated measurements was reported previously in children, and this may account for some of the differences that were seen between VSS and EXR measurements. This possibility emphasizes the importance of rechecking BPs with rigorous technique rather than relying on the more efficient VSS-type methods.

Because hypertension in the pediatric population is associated with significant target organ damage and morbidity, accurate BP screening and appropriate diagnostic evaluation is critical. The current recommendations are based on the careful use of aneroid devices, with appropriate cuff size on children who are relaxed and in a seated position for at least several minutes. Currently, if BP is measured at all, then the use of automated oscillometric devices at a VSS is a common practice that is efficient but inaccurate. Ideally, all pediatric groups would conform to task force recommendations, but at the very least, children with elevated VSS BPs should be reevaluated in the EXR by trained personnel. These results should not be interpreted as a reassurance that elevated BP that is detected in pediatric screening visits should be assumed to be normal. It is true that common screening practices will overestimate the BP; however, epidemiologic studies indicate that 2% to 4% of the pediatric population and between 15% and 30% of obese children have hypertension. Children need to be screened and assessed accurately for treatment to prevent significant long-term morbidity. An awareness that differences exist in routine oscillometric BP measurements that are taken on entrance to a VSS and aneroid BP values that are obtained in a quiet setting should decrease error in pediatric BP assessment.

ACKNOWLEDGMENT

This work was supported by National Institutes of Health grants DK-064587 and DK-071223 (to Dr Feig).

REFERENCES

Preventive Dental Care for Children in the United States: A National Perspective

Charlotte W. Lewis, MD, MPH\textsuperscript{a,b,c,d}, Brian D. Johnston, MD, MPH\textsuperscript{a,b,e}, Kristi A. Linsenmeyar, DDS\textsuperscript{d}, Alexis Williams, BS\textsuperscript{f}, Wendy Mouradian, MD\textsuperscript{a,b,d}

\textsuperscript{a}Division of General Pediatrics, Department of Pediatrics, \textsuperscript{b}Child Health Institute, and \textsuperscript{c}Department of Pediatric Dentistry, University of Washington, Seattle, Washington; \textsuperscript{d}Children’s Hospital and Regional Medical Center, Seattle, Washington; \textsuperscript{e}Department of Pediatrics, Harborview Medical Center, Seattle, Washington

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

ABSTRACT

OBJECTIVE. Preventive dental care is a cornerstone of optimal oral health. However, in 1996, only 38\% of US children received preventive dental care. We used the National Survey of Children’s Health to (1) describe the proportion of US children with $\geq$1 preventive dental visit within the previous year, (2) identify factors that were associated with preventive dental care use, and (3) test the hypothesis that preventive dental care use by near-poor children is associated with State Child Health Insurance Program policies for covering dental care.

METHODS. The National Survey of Children’s Health includes data from 102 353 children, weighted to represent 72.7 million children, nationally. Our outcome of interest was $\geq$1 preventive dental visit in the past year. We conducted multivariate regression analysis to identify factors that were associated significantly with this outcome using Stata survey capabilities.

RESULTS. In 2003, 72\% of US children had a reported preventive dental care visit in the previous year. On multivariable analysis, we found that being young, black or multiracial relative to white, lower income, and lacking a personal doctor were variables with a significantly lower likelihood of a preventive dental visit. Children in states with State Child Health Insurance Program dental coverage and broadest income eligibility had a 24\% higher likelihood of a preventive dental visit when compared with children in states with limited or no State Child Health Insurance Program coverage for dental services, on adjusted analysis.

CONCLUSIONS. Although the proportion of US children with a preventive dental visit now is higher than previously reported, children who are at highest risk for dental problems still are those who are least likely to receive preventive dental care. When states cover preventive dental care at income eligibility levels $\geq$200\% of the federal poverty level, there is a greater likelihood that near-poor children will receive preventive dental care.
PREVENTIVE DENTAL CARE is considered the cornerstone of optimal oral health promotion. Previous research from the 1996 Medical Expenditure Panel Survey (MEPS) indicated that 38% of US children had a preventive dental visit in the previous year, with even lower rates among younger, low-income, and minority children. With recent release of the National Survey of Children’s Health (NSCH), we have an important opportunity to describe and reassess preventive dental care use using a nationally representative pediatric data set.

In the past 10 years, a number of events with potential to increase access to preventive dental visits for children have occurred. Key among these was creation of Title XXI, the State Children’s Health Insurance Program (SCHIP), as part of the Balanced Budget Act of 1997. Under SCHIP, states could choose whether to receive federal matching funds to expand their Medicaid programs, purchase coverage through state-designed programs, or develop a combination of the 2 approaches.

With SCHIP enactment came the opportunity to expand health and dental care access for near-poor children who previously had exceeded income eligibility for Medicaid. However, SCHIP differs from Medicaid in a number of ways, including its coverage for dental care. Preventive and other dental care is a mandated benefit for Medicaid-eligible poor children through Early Periodic Screening, Diagnosis, and Treatment (EPSDT) legislation. In contrast, under SCHIP, states have the option of including dental care as part of covered services. At the onset of SCHIP’s implementation, 2 states, Colorado and Delaware, elected not to include dental care among SCHIP covered services. Florida provided SCHIP dental coverage only on a county-by-county basis. In 2003, Texas discontinued dental coverage under SCHIP (but recently restored it). The remaining states all cover preventive dental care and varying types of other dental treatment.

Under SCHIP, states also have more flexibility in establishing family income levels for eligibility. The maximum family income to qualify for SCHIP varies across states from 140% of the federal poverty level (FPL) to 350% of the FPL, whereas income eligibility for EPSDT-Medicaid is standard across all states; coverage is mandated for all eligible children who are younger than 6 years and have family incomes at or under 133% of the FPL and for children who are ≥6 and at or under 100% of the FPL, although certain states have extended Medicaid eligibility to higher incomes.

A number of reports have documented increases in the proportion of insured children and expanded access to health care since SCHIP’s implementation. Less is known about the effect of SCHIP on dental care use; the few published studies have focused on single states, where modestly positive effects were reported. In addition, a Mathematica Policy Research report on 27 states found that the implementation of SCHIP seems to have improved low-income children’s access to and use of dental services as of 2003 but there is variation across the states studied. We could identify no study to date that used national data in evaluating the impact of SCHIP and associated Medicaid expansions on preventive dental visits.

With this in mind, the objectives of this research were (1) to describe the proportion of US children in 2003–2004 with at least 1 preventive dental visit within the previous year, (2) to identify factors that were associated independently with preventive dental care use in the previous year, and 3) to test the hypothesis that preventive dental care use by near-poor children is associated with state SCHIP policies regarding eligibility for dental care.

METHODS

Data Source

We relied on data from the NSCH, which was sponsored and directed by the Maternal and Child Health Bureau of the US Health Resources and Services Administration along with the Centers for Disease Control and Prevention, and the National Center for Health Statistics. The survey was conducted between January 2003 and July 2004 and was intended to collect information on the physical, emotional and behavioral health and health care experience of a large sample of representative US children. The State and Local Area Integrated Telephone Survey (SLAITS) mechanism was used to complete 102,353 telephone surveys of households with children who were 0 to 17 years of age, with ~2000 respondents from each of the 50 US states and the District of Columbia. After identifying a household with children, the interviewers asked for the birth dates of all children who were younger than 18 years. One child in this age range then was randomly selected to be the subject of the interview. The respondent was the parent or guardian in the household who was most knowledgeable about the health and health care of the children who were younger than 18 years.

The data file is publicly available and contains population weights, stratum identifiers (ie, state name), and primary sampling unit codes that account for the complex sample design and permit population-based estimates with accurate SEs. Poststratification adjustments before release of the data ensure that population subgroups were represented properly in the weighted estimates that were generated from the data set.

Variable Selection and Sources

Our primary outcome of interest was report of a preventive dental visit during the previous 12 months, which
TABLE 1

Covariates That Were Hypothesized to Be Associated With a Preventive Dental Visit in the Previous Year

<table>
<thead>
<tr>
<th>Child</th>
<th>Parent</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dental insurance</td>
<td>Education</td>
<td>Ratio of practicing dentists per 100,000 population</td>
</tr>
<tr>
<td>Personal doctor&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Primary language</td>
<td>Percentage of dentists participating in SCHIP and in Medicaid</td>
</tr>
<tr>
<td>Age</td>
<td>Household income (% FPL)</td>
<td>Income eligibility level for SCHIP and program type</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td>Parent employment status</td>
<td>Medicaid/SCHIP child participation rates</td>
</tr>
<tr>
<td>Foreign born</td>
<td>MSA</td>
<td>Type of Medicaid/SCHIP Program</td>
</tr>
</tbody>
</table>

<sup>a</sup>The survey question inquired, “A personal doctor or nurse is a health professional who knows your child well and is familiar with your child’s health history. This can be a general doctor, a pediatrician, a specialist doctor, a nurse practitioner, or a physician assistant. Do you have 1 or more persons you think of as [child]’s personal doctor or nurse?”

was derived from this NSCH survey question, “During the past 12 months, did your child see a dentist for any routine preventive dental care, including checkups, screenings, and sealants?” This question was not asked of children who were younger than 12 months or who did not yet have any natural teeth. Before undertaking our data analysis, we identified potential variables that we hypothesized to be associated with preventive dental care use, on the basis of a comprehensive literature review. These variables included patient and family variables that were available in the NSCH. With regard to insurance, participants were asked whether they had health care insurance and whether this coverage was Medicaid or SCHIP (yes or no), without differentiating which of the 2 the child had. In addition, participants were asked whether the child had insurance to help pay for routine dental care but did not otherwise inquire about the specific kind of dental insurance.

We drew on other data sources to inform pertinent state variables, including (1) per capita distribution of general dentists, (2) proportion of dentists who participated in Medicaid and SCHIP, (3) percentage of Medicaid/SCHIP-eligible children who were enrolled, (4) maximum income by percentage of the FPL to qualify for Medicaid/SCHIP, and (5) type of Medicaid/SCHIP program (separate SCHIP, Medicaid expansion only, or combination).

To enhance our analyses, we created additional variables from those that were available in NSCH. We classified states by maximum income for SCHIP/Medicaid eligibility and dental coverage under SCHIP into the following categories:

1. States that did not offer statewide SCHIP dental services from program onset or that had discontinued dental coverage (Delaware, Colorado, Florida, and Texas): “no/limited SCHIP dental eligibility states.” These states all had separate SCHIP programs and annual Medicaid eligibility cutoffs equivalent to the minimum mandated (≤133% of the FPL for children 1–5 years and ≤100% of the FPL for ≥6 years).
2. States with SCHIP/Medicaid eligibility <200% of the FPL and that included dental coverage under SCHIP: “SCHIP dental/intermediate income eligibility states.”
3. States with SCHIP/Medicaid eligibility ≥200% of the FPL and that included dental services under SCHIP: “SCHIP dental/broadest income eligibility states.”

Within the last 2 state categories, we further stratified states into those with separate SCHIP or combination programs and those with only Medicaid expansion to reflect differences in how families may realize and perceive access to preventive dental care within the 2 systems. Stratification did not reveal any statistically significant differences in likelihood of preventive dental care, and we present only the nonstratified results here.

We developed a new race/ethnicity variable that combined the separate race and Hispanic ethnicity variables in the NSCH. Hispanic ethnicity combined with any race was classified as Hispanic, and the other races (white, black, multiracial, and other) were left unchanged. Using methods described by Mayer et al, we assigned a metropolitan statistical area (MSA) status to states where it had been suppressed to protect confidentiality in the original data. MSA was recoded from missing to non-MSA for states with small MSA samples (ie, most individuals lived in non-MSA or more rural locations [Alaska, Idaho, Maine, Montana, North Dakota, South Dakota, Vermont, and Wyoming]). Similarly, MSA status was recoded from missing to MSA for states with small non-MSA samples (Connecticut, Delaware, Hawaii, Massachusetts, Maryland, New Hampshire, Nevada, and Rhode Island). We also imputed poverty level for use in the multivariable regression models in the ~9000 cases for which this information was missing from the NSCH, using best subsets regression.

**Study Design and Data Analysis**

Data were analyzed with Stata 8.0 (Stata Corp, College Station, TX). To account for the complex survey design, we used Stata survey commands and the population
weights provided in the data files when generating population-level estimates and SE. We conducted descriptive analysis including bivariable tests of the association between explanatory covariates and our primary outcome: a preventive dental visit in the previous 12 months. In addition, we developed 2 multivariable logistic regression models. The first model included all of the hypothesized covariates of interest (Table 1) and was used to determine the independent associations between a preventive dental visit in the previous 12 months and the potential explanatory covariates.

The second model included an interaction term between the SCHIP eligibility dental coverage category and the child’s household income relative to the FPL. We relied on this model to test our hypothesis that near-poor children who reside in states with SCHIP dental coverage and broader income eligibility would have a higher likelihood of a preventive dental care visit. We estimated a linear combination of coefficients to determine the odds ratio (OR) for preventive dental care for the 3 categories of SCHIP dental and income eligibility (no/limited SCHIP dental eligibility states, SCHIP dental/intermediate income eligibility states, and SCHIP dental/broadest income eligibility states) for 3 groups of children as defined by EPSDT/Medicaid income eligibility criteria:

1. “Poor children” who would be Medicaid eligible only and not SCHIP eligible in any state. That is, their household income is ≤133% of the FPL if they are younger than 6 and ≤100% of the FPL for 6 and older. We defined this group to reflect differences in mandated Medicaid income eligibility by age, although technically it included some near-poor children who were 5 years and younger.

2. “Near-poor children,” that is, 1- to 5-year-old children at 133% to 199% of the FPL and children who are older than 5 years at 100% to 199% of the FPL. “Near-poor” children were of primary interest because they were the intended target of SCHIP/Medicaid expansions. This group also is where we would expect to see variation in preventive dental care use across states depending on the state’s SCHIP dental coverage and SCHIP/Medicaid income eligibility level.

3. “Higher income children” who had family incomes ≧400% of the FPL across all ages and would be neither SCHIP eligible nor Medicaid eligible in any state.

Assuming that we had accounted adequately for confounding variables, we posited that near-poor children who lived in the broadest SCHIP dental/broadest income eligibility states would have the greatest likelihood of preventive dental care in the previous 12 months and children who resided in states that had no/limited SCHIP dental eligibility would have the lowest. In contrast, we would not expect a trend across state SCHIP dental/ income eligibility categories in likelihood of preventive dental use among poor Medicaid-eligible (ie, non-SCHIP eligible) children regardless of state of residence because these children theoretically have consistent dental benefits across all states under EPSDT. Neither would we expect a trend across state dental and income eligibility categories for preventive dental visits among the higher income children, who presumably had other sources of coverage for dental care.

RESULTS

The 102,353 children on whom data were collected were weighted to represent 72.7 million children, nationally. Of these, 94% had responses to the preventive dental care question. Overall, 72% of US children were reported to have had a preventive dental visit in the previous 12 months. There was variability across states and the District of Columbia in the proportion of children with a preventive dental visit in the previous year. (Fig 1) Among children who were at or below 200% of the FPL, 62.5% had a reported preventive dental visit in the previous year. There were 6.35 million (8.6%) children who lacked health insurance and 16.3 million (22.8%) children who lacked dental insurance; the ratio of dental care uninsured to health care uninsured children was 2.6.

In examining the association of various potential explanatory variables on preventive dental care use, children who were 5 years and younger, were of nonwhite race/ethnicity, were of lower income, lacked dental insurance, lacked a personal doctor, lived in non-MSA areas or states, or were not born in the United States or whose parents were non–English speaking or unemployed or did not graduate from high school were significantly less likely, on bivariable analysis, to have had a preventive dental visit in the previous 12 months. (Table 2). Children who were 5 years and younger were the group with the lowest proportion receiving preventive dental care. There was variation on the state level (Fig 2), with Hawaii having the highest number of young children receiving preventive dental care (62%) and Nevada having the lowest (38%).

On multivariable analysis, being young, black or multiracial relative to white, having lower income, lacking dental insurance, lacking a personal doctor, living in a non-MSA area, being foreign born, and having parents with less than a high school education or who were unemployed all were independently associated with a statistically significant lower odds of a preventive dental visit in the previous year (Table 2). Being Hispanic and having a non-English primary language no longer were associated with a statistically significant lower likelihood of a preventive dental visit on the adjusted analysis.

Near-poor children in states with the SCHIP dental/broadest income eligibility had a 24% higher likelihood of a preventive dental visit relative to children in states with limited/no SCHIP dental services, on adjusted anal-
ysis. A statistically significant trend in ORs for preventive dental care among near-poor children was seen across SCHIP dental/income eligibility state categories (Fig 3). This trend was not seen among children who were outside of the near-poor group; that is, poor children who met Medicaid eligibility criteria (with theoretically consistent dental benefits across all states under EPSDT) and higher income children who would likely have other sources of payment for dental care regardless of SCHIP coverage within their state of residence. We did not find any significant difference in our results when we stratified states further by type of SCHIP program, whether Medicaid expansion, SCHIP, or combination.

DISCUSSION

We found that 72% of US children in 2003–2004 had at least 1 preventive dental visit in the previous year, as reported by their parent or other primary caregiver. Only a few previous studies separated out preventive dental visits from dental visits in general. Using 1996 MEPS data, Watson et al¹ found that 38% of children overall and 20% of children who were at or under 200% of the FPL had a preventive dental visit in the previous year. More recently, in a study of low-income children, Kenney et al⁴ found that 71% of children who were at or under 200% of the FPL had at least 1 reported preventive dental visit. In this current research, in which we relied on data from all 50 states and Washington, DC, as well as an order of magnitude larger sample size than these 2 previous studies, we found that 62.5% of low-income children had a reported preventive dental visit in the previous year. Certain groups continue to be underrepresented in their receipt of preventive dental care after adjustment for confounders. These include children who are young, who are black or multiracial, lack dental insurance, lack a personal doctor, are under 400% of the FPL (with a progressively negative impact as the household income drops), live in nonmetropolitan areas or states, or are foreign born.

Hispanic and white children were statistically similar in their likelihood of preventive dental care after controlling for other variables such as income, insurance, primary language, and being foreign born. Black children, however, have lower odds of preventive dental care that persists even after adjustment for income and insurance status. Among the many underlying factors of lower preventive dental care use in this population, cultural perceptions about preventive dental care and the lack of accessible professional dental care in predominantly black neighborhoods deserve additional attention.

We found an independent positive effect of having a personal doctor on the receipt of preventive dental care in the previous year. This is the first time to our knowledge that such an association has been recognized for the...
US pediatric population as a whole. We previously noted a similar relationship between having a personal doctor and not having an unmet dental care need among children with special health care needs. Families who identify a regular doctor for their child may have values and attributes that also lead them to seek preventive dental care; however, this relationship persisted despite controlling for a number of other confounders that also were associated with preventive habits and health care–seeking behaviors. It also is possible that physicians play an important role in referring patients to dental care or reinforcing the importance of preventive dental care to their patients and families. By whatever mechanism, there exists a link between having a personal doctor and receiving preventive dental care that deserves additional attention and study.

Younger children continue to be substantially under-

### TABLE 2

**Weighted Proportion of Children With a Preventive Dental Visit and Weighted Results of Bivariable and Multivariable Tests of Association for Outcome of a Preventive Dental Visit in the Previous Year (Weighted \( N = 72.7 \) Million)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Received Preventive Dental Visit in the Last Year, %</th>
<th>Adjusted OR (95% CI) From Multivariable Analysis&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, y</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–5</td>
<td>48.0</td>
<td>0.18 (0.17–0.19)</td>
</tr>
<tr>
<td>6–17</td>
<td>81.7</td>
<td>Reference</td>
</tr>
<tr>
<td>( P )</td>
<td></td>
<td>&lt;.0001&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>77.0</td>
<td>Reference</td>
</tr>
<tr>
<td>Hispanic</td>
<td>60.9</td>
<td>0.93 (0.82–1.06)</td>
</tr>
<tr>
<td>Black</td>
<td>66.5</td>
<td>0.69 (0.63–0.76)</td>
</tr>
<tr>
<td>Multiracial</td>
<td>68.2</td>
<td>0.71 (0.60–0.83)</td>
</tr>
<tr>
<td>Other</td>
<td>70.3</td>
<td>0.80 (0.65–0.99)</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Household income, % FPL</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \geq 400 )</td>
<td>82.4</td>
<td>Reference</td>
</tr>
<tr>
<td>300–399</td>
<td>80.1</td>
<td>0.84 (0.76–0.92)</td>
</tr>
<tr>
<td>200–299</td>
<td>74.3</td>
<td>0.67 (0.62–0.73)</td>
</tr>
<tr>
<td>100–199</td>
<td>65.8</td>
<td>0.48 (0.44–0.53)</td>
</tr>
<tr>
<td>&lt;100</td>
<td>58.2</td>
<td>0.44 (0.39–0.49)</td>
</tr>
<tr>
<td><strong>Household income, % FPL</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dental insurance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>77.2</td>
<td>Reference</td>
</tr>
<tr>
<td>None</td>
<td>57</td>
<td>0.41 (0.38–0.44)</td>
</tr>
<tr>
<td><strong>Dental insurance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Personal doctor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child has personal doctor</td>
<td>75.0</td>
<td>Reference</td>
</tr>
<tr>
<td>Child does not have personal doctor</td>
<td>57.7</td>
<td>0.59 (0.54–0.64)</td>
</tr>
<tr>
<td><strong>Personal doctor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MSA status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSA</td>
<td>72.4</td>
<td>Reference</td>
</tr>
<tr>
<td>Non-MSA</td>
<td>70.6</td>
<td>0.89 (0.83–0.95)</td>
</tr>
<tr>
<td><strong>Parent or respondent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Place of birth</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>72.8</td>
<td>Reference</td>
</tr>
<tr>
<td>Outside United States</td>
<td>59.7</td>
<td>0.75 (0.63–0.90)</td>
</tr>
<tr>
<td><strong>Place of birth</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English is primary language spoken in home</td>
<td>74.4</td>
<td>Reference</td>
</tr>
<tr>
<td>English is not primary language</td>
<td>55.7</td>
<td>0.91 (0.78–1.07)</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Parent or respondent with</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or higher education</td>
<td>73.6</td>
<td>Reference</td>
</tr>
<tr>
<td>Less than a high school education</td>
<td>54.1</td>
<td>0.79 (0.69–0.92)</td>
</tr>
<tr>
<td><strong>Parent or respondent with</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Employment status of parent or respondent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>73.2</td>
<td>Reference</td>
</tr>
<tr>
<td>Unemployed</td>
<td>63.0</td>
<td>0.88 (0.79–0.99)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Also adjusted for per-capita dentist distribution, the percentage of state dentists participating in Medicaid/SCHIP, and the percentage of Medicaid/SCHIP-eligible children who were enrolled.

<sup>b</sup> P values for bivariable test of association.
represented among those who received preventive dental care in the previous year. Children who were 1 to 5 years had an adjusted OR of a preventive dental visit of only 0.18 relative to children who were older than 5 years. There are a number of potential reasons for this. Despite recommendation by professional dental organizations, such as the American Academy of Pediatric Dentistry and the American Dental Association, for the first dental visit by 1 year of age, only 15% of pediatricians agreed with this concept in a 1999 national survey and therefore may not be recommending earlier dental visits for their patients. It was not until 2003 that the American Academy of Pediatrics specifically recommended an oral health evaluation by 1 year of age; before that, the

FIGURE 2
Preventive dental visits in children who were 5 years and younger, by state (more specific state-level results are available from the authors).

FIGURE 3
Adjusted OR of preventive dental care in the previous year for poor, near-poor, and higher income children by state SCHIP dental eligibility category. Adjusted for age, race, dental insurance, having a personal doctor, MSA status, parental education, English primary home language, and parental employment and state per capita dentist distribution, the percentage of state dentists participating in Medicaid/SCHIP, and the percentage of Medicaid/SCHIP-eligible children who were enrolled.

P trend = NS

Limited vs broadest: 24% increase

P trend = .03

SCHIP dental/broadest income eligibility states
SCHIP dental/intermediate income eligibility states
Limited or no SCHIP dental coverage states
recommended age was 3 years. Parents may not be aware of recommendations or may not perceive a need for dental care in their young children. In addition, general dentists, who make up the bulk of the dental workforce, may not feel comfortable caring for young children. Nevertheless, caries often have their onset in early childhood, particularly among those who are at high risk; National Health and Nutrition Examination Survey data that were collected between 1999 and 2002 indicated that 28% of US children who were between 2 and 5 years of age had caries, and substantially higher levels were observed among black, Mexican American, poor, and near-poor children. Therefore, it remains important to promote early initiation of preventive dental care for young children. Certain states perform substantially better than others in the proportion of children who are younger than 6 years and have had a preventive dental visit. This may reflect state and local efforts during the last 5 years to encourage preventive dental care use and increase access to professional dental care for very young children. These efforts largely grew out of the Surgeon General’s report and conference on children’s oral health in 2000.

We found that in states that cover preventive dental care at income eligibility levels at or >200% of the FPL through either SCHIP or Medicaid expansion, there is a greater likelihood of near-poor children’s receiving preventive dental care in that state. Under SCHIP, states have flexibility to establish income eligibility and whether dental care is a covered benefit. Thus, our findings bear consideration as we evaluate current SCHIP performance and contemplate potential revisions in anticipation of SCHIP legislative renewal in 2007. Certainly, improvement is needed; even in the best states, near-poor children lag substantially behind their higher income counterparts in preventive dental care use.

There are a few possible reasons for why we did not see a greater improvement in preventive dental use among children who were in the near-poor group and had the SCHIP dental coverage/broadest income eligibility. We controlled for a number of factors that would be expected to influence dental care access and use, such as the proportion of dentists within the state participating in Medicaid and SCHIP, as well as factors that reflect SCHIP outreach and accessibility using child SCHIP/Medicaid participation rates as a proxy. However, our adjustment may have been incomplete. In addition, 2003–2004 were challenging years for SCHIP. Many states faced budget crises that had an adverse impact on SCHIP and that led states to freeze or cap enrollment, implement copays or premiums, or reduce outreach efforts. Services under Medicaid largely were spared on the state level during these financially difficult times because of federal mandates that were in place. During this period, families may have heard of and been concerned about copayments, although these are prohibited for preventive care services under Title XXI legislation. Likewise, SCHIP-enrolled families may not have received adequate information about SCHIP dental coverage. This possibility was borne out in focus groups of SCHIP-enrolled parents, where some parents did not know that preventive dental care was a SCHIP covered service in their state.

Increasing dental insurance availability should be a priority. Although lack of dental insurance is 1 of a number of factors that contribute to disparities in preventive dental care use in the United States, it nevertheless is an important one, and it may be more amenable to intervention than are some of the other broader socioeconomic and racial/ethnic determinants. Increasing the proportion of children with dental insurance could be accomplished through SCHIP reform (ie, all states offer preventive dental services under SCHIP and expand income eligibility to include all near-poor children) and wider availability of employment-related dental benefits. To avoid crowd-out, states could offer a SCHIP dental-only benefits package that would be available to working families who received medical insurance but not dental insurance through their employer. This also would help to address the disparity in the ratio of dental-uninsured to health-uninsured children, which at 2.6 has remained constant since it last was reported using 1995 data.

Dental coverage is relatively inexpensive when compared with health insurance. In the late 1990s, actuarial estimates of the cost to provide comprehensive preventive, diagnostic, restorative, and select orthodontic care for all SCHIP-covered children ranged between $17 and $20 per child per month. More recently, a 2006 report for the Virginia State Proposed Model Insurance Product estimated that premium cost for private dental insurance for small business employees at 100% to 300% of the FPL would be $26/month for the employee and $75 for the employee and family.

Certain limitations bear mention. In this survey, preventive dental visits were determined by parental report. Parents may have overestimated the frequency of their children’s dental visits to provide more desirable responses. In addition, the phrasing of the NSCH preventive dental care question encompassed a broad spectrum of possibilities for preventive dental care; our results should not be taken to mean that children who had a preventive dental visit necessarily had a usual source of dental care or a “dental home.” Given the relatively limited number of variables that were available to us, there may be unmeasured variables that also had an impact on the likelihood of preventive dental care. For example, we had to rely on indirect measures of access to dental care (eg, proportion of dentists participating in Medicaid or SCHIP), and these may not reflect adequately some families’ experiences in seeking preventive dental care. We did not have data on in which public insurance program (Medicaid or SCHIP) in the child was...
enrolled. However, because we were interested in the overall effect of the state’s SCHIP dental coverage and SCHIP/Medicaid income eligibility on near-poor children as a group, it was less concerning to be lacking more specific child insurance information. Finally, because of the cross-sectional nature of this survey, establishing causality, for example, between having a personal doctor and receiving preventive dental care, was not possible.

Although this research was centered on preventive dental visits, the goals of such visits are to promote oral health and prevent disease. Unfortunately, caries concentrate in lower income children. Although this study identified higher rates of a preventive dental care visit than were measured using MEPS data in 1996, children who are at highest risk for dental problems, that is, lower income and minority children, still are those who are least likely to receive preventive dental care. Moreover, during early childhood, when establishing preventive habits and imparting preventive knowledge should be paramount, children are substantially less likely to have had a preventive dental visit; children who were younger than 6 years were only one fifth as likely to have received a preventive dental visit compared with older children.

With 62.5% of children at or below 200% of the FPL reported to have had a preventive dental visit, we now exceed the Healthy People 2010 goal of 57% of low-income children receiving preventive dental care in the previous year.33 Apart from asking whether this goal ever was good enough for poor and near-poor children, it was still a long way to go before we can say that all US children are receiving the preventive dental care that they need.

REFERENCES


e552 LEWIS et al


**Postterm Delivery and Risk for Epilepsy in Childhood**

Vera Ehrenstein, MPH\(^a\), Lars Pedersen, MSc\(^b\), Vibeke Holsteen, MD\(^c\), Helle Larsen, MD\(^d\), Kenneth J. Rothman, DrPH\(^a\), Henrik T. Sørensen, MD, PhD\(^a\)\(^b\)

\(^a\)Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts; \(^b\)Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; Departments of \(^c\)Paediatrics and \(^d\)Gynaecology and Obstetrics, Aalborg Hospital, Aalborg, Denmark

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

**ABSTRACT**

**OBJECTIVE.** Postterm delivery is a risk factor for perinatal complications, some of which increase risk for neurologic morbidity. We aimed to examine the association between postterm delivery and risk for epilepsy in childhood.

**METHODS.** We conducted a cohort study of singleton children who were born in 3 Danish counties from 1980 to 2001. Birth registry data were linked with hospital records to identify cases of epilepsy in the first 12 years of life. We included children who were born at ≥39 gestational weeks and computed crude, age-specific, and birth weight standardized incidence rates of epilepsy. We estimated adjusted incidence rate ratios according to mode of delivery by Poisson regression.

**RESULTS.** Among the 277 435 nonpreterm births, 32 557 were at ≥42 weeks, including 3396 at ≥43 weeks. Nearly one fourth of the 2805 epilepsy cases occurred in the first year of life. In that period, birth weight standardized incidence rate ratios for epilepsy were 1.3 for birth at 42 weeks and 2.0 for birth at ≥43 weeks, compared with birth at 39 to 41 weeks. Among children who were delivered by cesarean section, incidence rate ratios adjusted for birth weight, presentation, malformations, and county were 1.4 for birth at 42 completed weeks and 4.9 for birth at ≥43 weeks, compared with term vaginal births. There was a similar tendency among children who were delivered with the assistance of instruments. We found no evidence for the association between postterm delivery and risk for epilepsy beyond the first year of life.

**CONCLUSIONS.** Prolonged gestation is a risk factor for early epilepsy; the added increase in risk for instrument-assisted and cesarean deliveries could be attributable to factors that are related to both birth complications and epilepsy.
POSTTERM DELIVERY, DEFINED as delivery at or after 42 completed weeks (≥294 days) after the first day of the last menstrual period (LMP), occurs with a reported prevalence of 2% to 14% of deliveries. Other terms, such as “postdate,” and “postmature,” and “dysmature” sometimes have been used interchangeably with “postterm,” but “postterm” currently is preferred. Predictors of postterm delivery include anencephaly, hormonal disturbances, nulliparity, young maternal age, and history of prolonged pregnancy. The biological mechanisms that lead to postterm delivery are poorly understood and remain “a challenge for epidemiologic research.”

Children who are born postterm have higher perinatal mortality than term children. Population-based studies in Denmark and Sweden report a 25% increase in risks for stillbirth and neonatal death in infants who are born after 42 weeks of gestation, with mortality increase being even greater when postterm infants are growth restricted, meconium aspiration, asphyxia, infection, NICU admission. Reported neurologic complications include peripheral nerve paralysis, trauma of central nervous system (CNS), and convulsions. Previously, we and others found an association between depressed Apgar score and risk for childhood epilepsy; we noted that postterm delivery exacerbated the increase in risk.

Most studies of postterm delivery have focused on perinatal outcomes. Few studies have addressed long-term neurologic morbidity, and, to our knowledge, no formal epidemiologic study has examined the association between postterm delivery and risk for childhood epilepsy. Epilepsy is the most common serious neurologic disorder, with a heterogeneous and poorly understood etiology. The perinatal period is thought to play an important role in causing some cases of the disease. We examined the association between postterm delivery and the risk for epilepsy during the first 12 years of life.

METHODS
We used electronic medical databases to conduct a cohort study within the Danish counties of Aarhus, North Jutland, and Viborg. From the Danish National Birth Registry, we retrieved records of all live singleton births that occurred between January 1, 1980, and December 31, 2001. In Denmark, nearly all children with epilepsy are hospitalized. To identify cases of epilepsy, we used computerized local hospital discharge registries, which record, for each hospital admission, up to 20 discharge diagnoses and, starting in 1994, also contain data on outpatient and emergency visits. The discharge diagnoses were coded according to the International Classification of Diseases, Eighth Revision (ICD-8) through the end of 1993 and according to the ICD-10 thereafter. The codes for epilepsy were 345.xx in ICD-8 and G40.0 to G41.9 in ICD-10. Clinically, epilepsy is diagnosed according to the International League Against Epilepsy guidelines, which define epilepsy as either unprovoked seizure episodes, relevant electroencephalography findings, or both. We defined the outcome of epilepsy as the first-time hospitalization or outpatient visit recorded in the hospital discharge registry with a diagnosis of epilepsy within the first 12 years of life. The date of diagnosis was considered to be the date of epilepsy onset.

Data on emigration and death were retrieved from the Danish Civil Registration System, which after being created in 1968, covers the entire Danish population and is updated daily. We linked records from the different registries by the civil registration number, which is a unique administrative identifier that is assigned at birth to Danish residents.

In the birth registry, gestational age was recorded in completed weeks through 1996 and in fractional weeks (based on days) thereafter. We defined postterm delivery, according to the World Health Organization guidelines, as delivery at ≥42 completed weeks of gestation. We calculated incidence rates of epilepsy and 95% confidence intervals (CI) by gestational age. For the main analysis, we included nonpreterm (≥39 weeks) births and computed crude and birth weight standardized incidence rates of epilepsy (using the person-time distribution in 500-g birth weight categories of the study cohort as the standard). We then calculated incidence rates of epilepsy in categories of completed gestational weeks (39-41, 42, and ≥43) according to the recorded age of onset. Within each gestational age category, we used Poisson regression to obtain fitted values of age-specific incidence rates, with age in completed years, age squared, and the square root of age as predictors. We created an indicator variable for each combination of mode of delivery (unassisted vaginal, forceps/vacuum-assisted, and cesarean) and gestational age category and used them in a Poisson model to examine variation of the gestational age-epilepsy incidence rate ratio (IRR) by mode of delivery.

Because idiopathic epilepsies tend to have genetic determinants and therefore may be unrelated to prolonged gestation, we did a subanalysis restricting the case definition to epilepsy cases that were not listed as idiopathic (censored when ICD-10 codes were G40.0 or G40.3). We did this subanalysis in the cohort of children who were born in 1994-2001, because their diagnoses were recorded using the ICD-10, in which idiopathic epilepsies are listed specifically.

There was little confounding by any of the variables.
that were available for analysis. For the adjusted analyses, we retained fetal presentation as a potential risk factor for neurologic morbidity, county of birth as a marker for coding practices, and an indicator for birth defects diagnosed at birth because having such a birth defect was a strong predictor of epilepsy in our data and affected the magnitude and the precision of IRR estimates for forceps-assisted deliveries. (In addition, a recent study found maternal use of antiepileptic medication to be a risk factor for birth defects, making the latter a correlate of maternal epilepsy.29) Birth weight may be a correlate of traits that causally link prolonged gestation and risk for epilepsy (eg, malnutrition in overly small infants, trauma as a result of large size in overly heavy infants), which is an argument against using birth weight for adjustment in this analysis. Removing birth weight from any of the analyses had little effect on the results. We retained birth weight for the standardization and adjustment, because it may be a marker of antenatal hormonal or metabolic perturbations that may affect both the size and the susceptibility to disease.29 Variables for infant’s gender and birth order; confidence of the LMP; and mother’s age, smoking, and cohabiting with the partner were not used in the final analytic models because in these data they were not associated with epilepsy risk and their inclusion did not change the results. We analyzed the data with SAS 9.01 software (SAS Institute, Cary, NC) and used Episheet30 to calculate standardized estimates.

RESULTS

During the study period, 338 633 single live births were recorded in the 3 counties. We excluded 336 (0.1%) records with missing gestational age. Among the remainder, 277 435 (82%) births occurred at 39 completed weeks of gestation or later. Of these, 32 557 (12%) infants were delivered postterm, including 3396 (1%) delivered at 43 weeks or later.

Compared with term newborns, postterm infants were less likely to weigh ≤2500 g and more likely weigh >4000 g, to undergo cesarean or instrument delivery, to be first born, or to have a 5-minute Apgar score <7. Compared with mothers who delivered at term, mothers who delivered postterm were somewhat more likely to have smoked during pregnancy (Table 1).

Across the full range of gestational age, the crude incidence rate of epilepsy decreased with increasing gestational age through week 41, after which it increased again, although the precision of estimates was low at both extremes of gestational age distribution (Fig 1). Overall, 2805 epilepsy cases were recorded among the members of the cohort during the follow-up period. Regardless of gestational age, incidence rates of epilepsy were the greatest in the first year of life, with 657 (23%) of the cases recorded with the onset during this period (Table 2; Fig 2). In year 2 after birth, there were 357 (13%) cases; in year 3, there were 283 (10%); in year 4, there were 240 (8%); in year 5, there were 197 (7%); in

### Table 1: Newborn and Maternal Characteristics and Postterm Delivery

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Gestational Age, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>39–41 wk (N = 244 878)</td>
</tr>
<tr>
<td>Birth weight, g</td>
<td>≤2500</td>
</tr>
<tr>
<td></td>
<td>2501–3000</td>
</tr>
<tr>
<td></td>
<td>3001–3500</td>
</tr>
<tr>
<td></td>
<td>3501–4000</td>
</tr>
<tr>
<td></td>
<td>4001–4500</td>
</tr>
<tr>
<td></td>
<td>4501–5000</td>
</tr>
<tr>
<td></td>
<td>&gt;5000</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
</tr>
<tr>
<td>Mode of delivery</td>
<td>Unassisted vaginal</td>
</tr>
<tr>
<td></td>
<td>Vacuum assisted</td>
</tr>
<tr>
<td></td>
<td>Forceps assisted</td>
</tr>
<tr>
<td></td>
<td>Cesarean</td>
</tr>
<tr>
<td>Fetal presentation</td>
<td>Cephalic</td>
</tr>
<tr>
<td></td>
<td>Breech</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
<tr>
<td></td>
<td>Missing</td>
</tr>
<tr>
<td></td>
<td>Boy</td>
</tr>
<tr>
<td></td>
<td>Apgar score &lt;7 at 5 min</td>
</tr>
<tr>
<td></td>
<td>Birth defect discovered at birth</td>
</tr>
<tr>
<td></td>
<td>First born</td>
</tr>
<tr>
<td></td>
<td>Mother’s age at delivery, y</td>
</tr>
<tr>
<td>&lt;20</td>
<td></td>
</tr>
<tr>
<td>21–25</td>
<td></td>
</tr>
<tr>
<td>26–30</td>
<td></td>
</tr>
<tr>
<td>31–35</td>
<td></td>
</tr>
<tr>
<td>36–40</td>
<td></td>
</tr>
<tr>
<td>≥40</td>
<td></td>
</tr>
<tr>
<td>Mother is unsure of LMP start</td>
<td>30 567 (12)</td>
</tr>
<tr>
<td>Mother smoked during pregnancy*</td>
<td>33 057 (25)</td>
</tr>
<tr>
<td>Mother not living with a partner*</td>
<td>10 438 (8)</td>
</tr>
</tbody>
</table>

* Data available on children who were born after 1990 (N = 147 656).

FIGURE 1

Crude incidence rate of epilepsy (cases per 10 000 person-years), according to completed gestation, with pointwise 95% Poisson CIs.
year 6, there were 204 (7%); in year 7, there were 170 (6%); in year 8, there were 163 (6%); in year 9, there were 157 (6%); in year 10, there were 158 (6%); in year 11, there were 117 (4%); and in year 12, there were 102 (4%). The first year of life was the period when the differences in the epilepsy rates according to gestational age were most pronounced (Fig 2). We therefore focused subsequent analysis on epilepsy that occurred before age 1.

Of the epilepsy cases with the recorded onset during the first year of life, the proportion of girls with the diagnosis was ~48% regardless of gestational age at birth. Table 3 shows crude and birth weight standardized incidence rates and rate ratios. The standardized IRR was 1.3 (95% CI: 1.0–1.7) for 42 completed weeks of gestation and 2.0 (95% CI: 1.2–3.5) for gestation of 43 weeks or longer.

The magnitude of the IRR varied according to delivery mode (Table 4). For children who were delivered vaginally without help of vacuum or forceps, postterm gestation of any length was associated with a 1.3-fold increase in epilepsy risk in the first year of life. Among children who were born by cesarean section, delivery at 42 weeks was not related to an increase in risk beyond that associated with cesarean section itself, but delivery at 43 weeks or later was associated with a 3.5-fold increase in risk compared with term cesarean births and a nearly fivefold increase in risk compared with term unassisted vaginal births. For vacuum- or forceps-assisted deliveries, our point estimates showed a dose–response pattern, with longer postterm gestation conferring greater epilepsy risk. When mode of delivery was added as a covariate to the other variables that were used in the adjusted analysis, adjusted IRR was 1.3 (95% CI: 1.0–1.6) for birth at 42 weeks and 1.9 (95% CI: 1.1–3.1) for birth at 43 weeks onward. In this analysis, adjusted IRR were 1.4 (95% CI: 1.0–1.8) for cesarean delivery and 1.1 (95% CI: 0.8–1.4) for vacuum/forceps-assisted delivery, compared with unassisted vaginal delivery.

Of the 202 epilepsy cases recorded in the first year of life among children who were born in 1994–2001, 36 (18%) were recorded as idiopathic. Rates of epilepsy excluding these cases (per 10 000) were 15 (95% CI: 12–17) for children who were born during weeks 39 to 41 of gestation and 23 (95% CI: 15–33) for children who were born at 42 weeks or later; crude rate ratio was 1.6 (95% CI: 1.0–2.3). An analysis of a “low-risk” subgroup (restricted to children with optimal birth weight for gestational age [between 10th and 90th percentiles], no malformation at birth, Apgar score >6 at 5 minutes, and vaginal delivery in cephalic presentation without assistance of forceps or vacuum) yielded IRR of 1.4 (95% CI: 1.0–1.8).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Gestational Age, Completed wk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>39–41 (n = 244 848)</td>
</tr>
<tr>
<td>Cases</td>
<td>2429</td>
</tr>
<tr>
<td>Person-years</td>
<td>2 311 920</td>
</tr>
<tr>
<td>Crude incidence rate, per 10 000 person-years</td>
<td>10</td>
</tr>
<tr>
<td>Crude IRR (95% CI)</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Epilepsy diagnosed during the first year of life

| Cases     | 560 | 82 | 15 | 657 |
| Person-years | 243 329 | 28 959 | 3368 | 275 656 |
| Crude incidence rate, per 10 000 person-years | 23 | 28 | 44 | 24 |
| Crude IRR (95% CI) | 1.0 | 1.2 (1.0–1.5) | 1.9 (1.1–3.2) | |
for delivery at 42 completed weeks and 1.7 (95% CI: 0.6–3.7) for delivery at 43 weeks or later.

We were able to access data on labor induction for the small subset of the most recent (1997–2001) births in our study cohort. Restricting this subcohort to vaginal deliveries, we calculated incidence rates of epilepsy in the first year of life, according to labor induction and postterm delivery (Table 5). Although all of the CIs overlap and the data are sparse, the point estimates suggest that the effect of prolonged gestation on the risk for epilepsy may be more pronounced in the absence of labor induction.

**DISCUSSION**

In this large, population-based cohort study, we found delivery at ≥42 weeks of gestation to be associated with an increased risk for epilepsy in the first year of life. The magnitude of the risk increase depended on the duration of prolonged gestation. Little evidence for an association in subsequent years lends support to the conjecture that perinatal causes and postterm delivery in particular play a greater role in determining early-life neurologic susceptibility. During the first year of life 42 (6%) of 657 children with epilepsy also had a diagnosis of cerebral palsy: 36 among term infants (incidence: 1.5 in 10 000) and 6 among infants who were born at 42 completed weeks (incidence: 2.0 in 10 000). These findings suggest that prolonged gestation may be a risk factor for other neurologic disability.

The overall prevalence of postterm deliveries did not change substantially with the year of birth in our cohort, but the prevalence of deliveries at 43 weeks or more decreased from ~2% in the 1980s to <0.5% in 2000–2001. Thus, over time, very late deliveries have accounted for a decreasing proportion of postterm births. This could be caused either by a true decrease of prevalence resulting from more routine labor induction before 43 weeks or by better pregnancy dating. We calculated IRRs for epilepsy in the first year of life in 3 strata of birth year (with comparable prevalence of delivery at 43+ weeks): 1980–1990, 1991–1993, and 1994–2001. The respective periods’ IRRs were 1.0 (95% CI: 0.7–1.5), 1.3 (95% CI: 0.8–2.0), and 1.5 (95% CI: 1.0–2.2) for delivery at 42 completed weeks and 1.8 (95% CI: 0.9–3.2), 1.9 (95% CI: 0.5–5.1), and 2.0 (0.3–6.4) for delivery at 43 completed weeks or later. The slight upward trend of the point estimates is consistent with the conjecture of

**TABLE 3**  
Occurrence of Epilepsy in the First Year of Life by Birth Weight and Gestational Age  

<table>
<thead>
<tr>
<th>Parameter</th>
<th>39–41 Completed wk Gestation</th>
<th>42 Completed wk Gestation</th>
<th>≥43 Completed wk Gestation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth weight, g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤2500</td>
<td>17/2381 (71)</td>
<td>0/92 (0)</td>
<td>1/15 (667)</td>
</tr>
<tr>
<td>2501–3000</td>
<td>65/23 351 (28)</td>
<td>9/1443 (62)</td>
<td>1/173 (58)</td>
</tr>
<tr>
<td>3001–3500</td>
<td>186/83914 (22)</td>
<td>21/6807 (31)</td>
<td>2/771 (26)</td>
</tr>
<tr>
<td>3501–4000</td>
<td>195/90 836 (21)</td>
<td>28/11 499 (24)</td>
<td>6/1312 (46)</td>
</tr>
<tr>
<td>4001–4500</td>
<td>73/32 563 (21)</td>
<td>21/6935 (30)</td>
<td>4/777 (51)</td>
</tr>
<tr>
<td>4501–5000</td>
<td>16/6325 (25)</td>
<td>2/1848 (11)</td>
<td>1/278 (36)</td>
</tr>
<tr>
<td>&gt;5000</td>
<td>2/804 (25)</td>
<td>1/279 (36)</td>
<td>0/40 (0)</td>
</tr>
<tr>
<td>Standardized incidence rate (per 10 000)</td>
<td>22</td>
<td>30</td>
<td>46</td>
</tr>
<tr>
<td>Standardized IRR (95% CI)</td>
<td>1.0 (1.0–1.7)</td>
<td>2.0 (1.2–3.5)</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 4**  
Postterm Delivery and Epilepsy in the First Year of Life: Adjusted IRRs (95% CIs), Stratified by Mode of Delivery  

<table>
<thead>
<tr>
<th>Mode of delivery</th>
<th>39–41 wk</th>
<th>42 wk</th>
<th>≥43 wk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>IRR (95% CI)</td>
<td>Cases</td>
</tr>
<tr>
<td>Unassisted vaginal</td>
<td>452</td>
<td>1.0</td>
<td>63</td>
</tr>
<tr>
<td>Vacuum/forceps assisted</td>
<td>50</td>
<td>1.1 (0.8–1.5)</td>
<td>10</td>
</tr>
<tr>
<td>Caesarean</td>
<td>58</td>
<td>1.3 (1.0–1.8)</td>
<td>9</td>
</tr>
</tbody>
</table>

IRRs Adjusted for birth weight, county of birth, fetal presentation, and presence of birth defects detected at birth.

**TABLE 5**  
Incidence Rates of Epilepsy in the First Year of Life, According to Labor Induction and Gestational Age, Among Children Who Were Born Vaginally in 1997–2001  

<table>
<thead>
<tr>
<th>Vaginal Delivery</th>
<th>Births</th>
<th>Cases</th>
<th>Incidence Rate (95% CI), Per 10 000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noninduced</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Term</td>
<td>45 696</td>
<td>66</td>
<td>14 (11–18)</td>
</tr>
<tr>
<td>Postterm</td>
<td>3776</td>
<td>8</td>
<td>21 (10–42)</td>
</tr>
<tr>
<td>Induced</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Term</td>
<td>2747</td>
<td>6</td>
<td>22 (10–48)</td>
</tr>
<tr>
<td>Postterm</td>
<td>1763</td>
<td>3</td>
<td>17 (&lt;1–50)</td>
</tr>
</tbody>
</table>
better pregnancy dating, resulting in diminished misclassification of gestational age over time.

Postterm delivery may precipitate or accelerate epilepsy through increasing risk for infectious, hypoxic, or mechanical injury to the developing brain. A potential causal mechanism for the observed association could involve intrauterine exposure to meconium, occurring in up to 30% of postterm births. Meconium passage is associated with impaired fetal oxygenation and may increase risk for fetal bacterial invasion, leading to inflammatory brain damage. Prolonged pregnancy also is associated with up to a 33% decrease in amniotic fluid volume, which can lead to fetal distress and cause early CNS damage.

Increased likelihood of labor induction and instrument deliveries after prolonged pregnancy, coupled with greater fetal size, set the stage for mechanical injuries during labor and delivery. Approximately 24% of deliveries that occur after the 42nd week of gestation in Denmark are induced. Induced contractions may be forceful and prolonged, causing extended occlusion of the umbilical cord, reduced oxygenation, and fetal acidosis, all of which may contribute to neurologic damage. Labor induction could be a confounder in this study if contraindications for labor induction, such as breech presentation or maternal severe hypertension, also were markers of a fetus’s neurologic susceptibility. At the same time, the analysis of the recent births with available induction data suggests that, at least for vaginal deliveries, the effect of prolonged gestation on the risk for epilepsy may be more pronounced when the labor is not induced.

In a study of singleton infants who weighed 2500 to 4000 g and were born in cephalic presentation to nulliparous mothers, Towner et al found that birth that was assisted by vacuum or forceps, compared with unassisted vaginal birth, was associated with up to a threefold increased risk for cerebral hemorrhage and CNS depression and nearly a twofold increased risk for convulsions; the increase in risk was doubled or tripled when a failed vacuum extraction necessitated the use of forceps (which occurs in 9% to 14% of vacuum-assisted deliveries). In our cohort, instrument delivery at term alone was associated with only a slight increase in epilepsy risk, as was postterm delivery alone (Table 4). The 2 conditions together, however, were more likely to produce a more substantial increase in risk. This observation also is consistent with the current view that some cases of epilepsy are a result of accumulated CNS insults. A limitation of our study is the inability to distinguish between emergency and elective cesarean sections and between forceps deliveries with and without attempted vacuum extraction.

Although, until now, no study specifically assessed the association of prolonged gestation with epilepsy risk, large population-based studies in Nordic countries report an association between birth after week 42 and increased incidence of convulsions, which are symptoms of epilepsy. A Danish study of ∼110 000 nonpreterm births found an adjusted relative risk for convulsions to be 1.4 (95% CI: 0.90–2.1), and in a Swedish study of >500 000 nonpreterm births, the relative risk estimate for convulsions was 1.5 (95% CI: 1.2–2.0). These estimates are similar to the birth weight standardized IRRs for epilepsy in our study.

Lack of detailed clinical data is a limitation of our study. For example, because ICD-8, which was used to code diagnoses through 1993, does not have a special code for infantile spasms, it is possible that this common early-life epilepsy type was underascertained for children who were born before 1994. (ICD-10, used from 1994 onward, lists infantile spasms under the heading G40.4, which was included in our case definition.) We noted that more than half of the epilepsy cases in the first year of life received a code of “unspecified” or “other,” precluding valid inferences about distribution of epilepsy types on the basis of routine registration data. At the same time, the abbreviated nature of computerized records may be only partially to blame for this drawback. As Korff and Nordli noted in their 2006 review, many infants with epilepsy “do not fit in any of the currently used subcategories.”

The validity of our results depends on accurate classification of gestational age. Postterm deliveries accounted for 9.1% of all births in our cohort. As reviewed by Shea et al, studies that used combined early ultrasound/LMP dating method place the prevalence of postterm delivery between 4% and 7%. From published tables of a Danish nationwide study, we back-calculated the prevalence of postterm deliveries in 1978–1993 to be ∼7.6% of all births. Reported prevalence depends to a large extent on the definition of postterm (as many as 8 different definitions are cited in studies). In our data, excluding children who were born with gestational age recorded at exactly 42 weeks reduced the prevalence to 2.3% of all births. We included week 42 in our definition of postterm in accordance with current guidelines. Increased risk for some outcomes has been reported for births as early as week 41 of gestation. The present definition, although undoubtedly subject to misclassification, is a compromise between over- and underascertainment of true postterm deliveries.

Prevalence of postterm delivery also depends on the method of gestational age determination and on variations in menstrual cycle length (the LMP-based gestational age dating assumes a 28-day menstrual cycle). Midwives in Denmark record gestational age in completed weeks after LMP and, whenever necessary, make corrections after ultrasound examination. The use of ultrasound in Denmark increased during our study period. Nevertheless, the prevalence of postterm births in our cohort decreased only slightly over time, suggesting...
that increased use of ultrasound did not affect pregnancy dating profoundly, at least at the upper extreme of gestational age distribution. Although a good general agreement has been reported between LMP- and ultrasound-based gestation estimation,\textsuperscript{42} misclassification tends to occur at the extremes of the distribution. According to some studies, LMP-based dating tends to overestimate the prevalence of postterm delivery by \( \sim 9\% \),\textsuperscript{43} whereas ultrasound dating, in some cases, may underestimate the relative risk for postterm delivery by \( > 10\% \).\textsuperscript{44}

Nondifferential misclassification of gestational age would be expected to cause attenuation of relative effects.\textsuperscript{45} Also, the incidence rate of epilepsy in this study is slightly higher than other reports from developed countries,\textsuperscript{19} which may reflect overascertainment of epilepsy in electronic discharge records.\textsuperscript{13} Any errors regarding the ascertainment of epilepsy are likely to be independent of the child’s gestational age and therefore also would result in underestimation of the effect. The relative effect could be overestimated if misclassification of febrile seizures as epilepsy were more likely to occur among children who are born postterm. At the same time, children with febrile seizures are more likely to develop some types of epilepsy later in life.\textsuperscript{46}

CONCLUSION
We offer evidence that prolonged gestation is a risk factor for epilepsy in the first year of life.

ACKNOWLEDGMENTS
This study was funded by the Western Danish Research Forum for Health Sciences. Ms Ehrenstein is supported by the Ruth L. Kirschstein Individual Predoctoral Fellowship from the National Institute of Neurological Disorders and Stroke.

We thank the reviewers for the excellent comments that helped improve this article.

REFERENCES
31. Ahanya SN, Lakshmanan J, Morgan BL, Ross MG. Meconium
Comparison of Current Health, Functional Limitations, and Health Care Use of Young Adults Who Were Born With Extremely Low Birth Weight and Normal Birth Weight

Saroj Saigal, MD, FRCP(C)a, Barbara Stoskopf, RN, MHScsa, Michael Boyle, PhDb, Nigel Paneth, MD, MPHc, Janet Pinelli, RNC, MScN, DNSa, d, David Streiner, PhDde, John Goddeeris, PhDf

Departments of aPediatrics and bPsychiatry and Neurosciences and dSchool of Nursing, McMaster University, Hamilton, Ontario, Canada; Departments of cEpidemiology, Pediatrics, and Human Development and eEconomics, Michigan State University, East Lansing, Michigan; fDepartment of Psychiatry, University of Toronto, Toronto, Ontario, Canada

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

**ABSTRACT**

OBJECTIVE. The objective of this study was to compare the current health status, physical ability, functional limitations, and health care use of extremely low birth weight and normal birth weight young adults.

METHODS. A longitudinal study was conducted of a population-based cohort of 166 extremely low birth weight survivors (501–1000 g birth weight; 1977–1982 births) and a group of 145 sociodemographically comparable normal birth weight individuals. Current health status, history of illnesses, hospitalizations, use of health resources, and physical self-efficacy were assessed through questionnaires that were administered to the young adults by masked interviewers.

RESULTS. Individuals completed the assessments at a mean age of 23 years. Neurosensory impairments were identified in 27% of extremely low birth weight and 2% of normal birth weight individuals. No differences were reported in the current health status for physical or mental summary scores. Extremely low birth weight young adults reported a higher prevalence of chronic health conditions in the past 6 months. A significantly higher proportion of extremely low birth weight individuals had functional limitations in seeing, hearing, and dexterity and experienced clumsiness and learning difficulties. Except for prescription glasses, medications for depression, and home-care services for extremely low birth weight individuals, there were no significant differences between groups in use of health care resources. Extremely low birth weight individuals had significantly weaker hand grip strength and lower scores for physical self-efficacy, perceived physical ability, and physical self-confidence.

CONCLUSIONS. Extremely low birth weight young adults seem to enjoy similar current health status to their normal birth weight peers. However, they continue to have...
significantly poorer physical abilities and a higher prevalence of chronic health conditions and functional limitations. Contrary to expectations, they do not pose a significant burden to the health care system at young adulthood.

Recent reports on the long-term outcomes of very low birth weight (VLBW) and extremely low birth weight (ELBW) infants who were born in the early post-neonatal intensive care era have moved beyond midchildhood and adolescence into young adulthood. At adolescence, most reports indicate that substantial morbidity persists in intellectual status, school attainment, 1–3 behavioral difficulties, 4 and lower growth attainment compared with the normal birth weight (NBW) group. 5–7 Although there is some reduction in acute morbidity compared with the normal birth weight (NBW) group,1–3 behavioral difficulties, 4 and lower growth attainment compared with the normal birth weight (NBW) group,5–7 some reduction in acute morbidity persists in intellectual status, school achievement, 1–3 behavioral difficulties, and lower growth attainment compared with the normal birth weight (NBW) group.5–7 Although there is some reduction in acute morbidity, the VLBW individuals have significantly higher rates of functional limitations, greater compensatory dependence, and increased use of health care resources.8,9

In the few studies that have pursued additional follow-up, many of the educational and growth disadvantages that are associated with being of ELBW and VLBW have persisted to adulthood.10–16 However, little is known of the functional limitations, health status, and health care needs of these vulnerable young adults (YAs) once they are too old to use pediatric services. Such information is necessary to project and plan for medical services beyond those that routinely are required by the general population at adulthood and middle age.

In this report, we present data at young adulthood on the general health, physical abilities, functional limitations, and health care use of a regional cohort of former ELBW infants who have been followed longitudinally from birth17 in comparison with a term-born NBW group.18 We hypothesized that although ELBW YAs would continue to have more chronic health problems, greater functional limitations, and poorer physical abilities than NBW YAs, there would be no differences in their current health status; furthermore, the absolute rates of use of health care services would decline even further than previously reported at adolescence.6

Methods
Participants
ELBW survivors (501–1000 g birth weight) who were born between 1977 and 1982 to residents of a geographically defined region in central-west Ontario were followed longitudinally from birth.17 We previously reported on their growth and health outcomes at age 8,18 in adolescence,6 and at young adulthood.16

At 8 years of age, term-born children (1977–1981 births) who were comparable in gender, age, and social class19 to the ELBW group were randomly recruited from a list that was provided by the local school boards and followed longitudinally.18

Interview Protocol
The YAs were the primary respondents for all questionnaires and were interviewed by lay professional interviewers who were unaware of the group status. The interview process involved a structured format using scripted questions with a skip pattern that was administered in the same order. The majority (93%) of interviews were conducted at McMaster Children’s Hospital (Hamilton, Ontario, Canada) between January 1, 2002, and April 30, 2004. Information regarding health status and health care use of YAs with severe impairment was obtained from their parents.

Questionnaires
Ethics approval was obtained from the Research Ethics Board of Hamilton Health Sciences, and written informed consent was obtained from all YAs and their parents.

Demographics
Age, marital status, current living arrangements, household membership, educational attainment and employment were obtained from the YAs by direct interview. Parents provided information on their own education and current employment. Both maternal and paternal variables were included to assign parental socioeconomic status.19

General Health Information
The following standardized questionnaires were used:

- SF-3620: A widely used and well-validated 36-item questionnaire measures physical and mental health in the previous 4 weeks across 8 domains. Physical health summary score is derived from 4 subscales: physical functioning (10 items that describe ability to do physical tasks), role—physical (4 items about difficulties or limitations in physical ability), bodily pain (2 items), and general health (5 items). Mental health summary score also has 4 subscales: vitality (4 items that measure energy level), social functioning (2 items regarding physical and mental health interfering with normal social life), role—emotional (3 items about effect of mental problems on everyday life), and mental health (5 items that measure mood). Scores were converted to norm-based scores (mean: 50; SD: 10). In addition, respondents were asked to rate their health in 1 of the following categories: excellent, very good, good, fair, or poor.

- Ontario Child Health Study Questionnaire21: Provides information on chronic health conditions, functional limitations, and emotional and mental health issues (diagnosed by a health professional and lasting 6
months or longer); absenteeism at work or at school as a result of chronic health problems; and limitations in normal activities as a result of health problems.

- Canadian Community Health Survey (Statistics Canada)\(^22\)\(^-\)\(^24\): Data from cycles 2.1 (2003) and 3.1 (2005) on the health of YAs (aged 20–24) are presented as footnotes in the tables where relevant information is available.

**Physical Self-Efficacy Scale**
The Physical Self-Efficacy Scale\(^25\) is a self-administered questionnaire that provides information on total physical self-efficacy, perceived physical ability, and physical self-presentation confidence.

**Hand-Grip Strength**
Hand-grip strength was measured independently in both hands using a dynamometer. Data are presented for dominant hand for participants without neurosensory impairments (NSI).

**Use of Health Care Resources**
The questionnaire was based on several sources.\(^26\),\(^27\) The information included visits to all health professionals, outpatient tests, and use of home-care services in the past 6 months and hospitalizations and surgery in the past 12 months. All medications that were taken on a regular basis (and reasons for taking) were noted. Use of mechanical aids or assistive devices and home/car/house adaptations, etc, were also recorded.

**Statistical Analyses**
\(\chi^2\) tests of significance were used to test differences in categorical variables between groups (ELBW versus NBW) and gender-specific differences. Fisher’s exact test was used when necessary. For variables with significant differences, analysis of variance to compare mean differences between groups and odds ratios (ORs) and 95% confidence intervals (CI) were calculated. \(T\) tests were used to compare mean values. Although exact \(P\) values are provided where applicable, because of multiple testing, Holm’s correction\(^28\) was applied to all \(P < .05\) separately for each table (for total group and by gender) to establish statistical significance. Values that were found to be significant by Holm’s correction are indicated with a superscript Holm’s beside the \(P\) values. SPSS11.0 (SPSS, Chicago, IL) was used for all statistical analyses.

**RESULTS**

**Study Participants**

**ELBW Participants**
Of 397 livebirths, 179 (45\%) survived to hospital discharge\(^17\); 13 children subsequently died: 6 before 3 years of age, 4 with severe NSI between 9 and 16 years, and 3 in the late teens (1 NSI). Of 166 individuals available, 9 were lost and 8 refused (6 of these 17 had NSI). The outcome is reported on 149 (90\%) of 166 ELBW YAs, including 7 with severe NSI for whom parental proxy responses were obtained.

**NBW Participants**
Of the 145 term control subjects,\(^18\) 5 were lost to follow-up, 7 refused (none had NSI), and the remaining 133 (92\%) participated.

**Birth Characteristics and Sociodemographics of Parents and YAs**
Mean birth weight for the ELBW cohort was 841 g (SD: 124 g), and mean gestational age was 27.1 (SD: 2.3) weeks (Table 1). More than one quarter of ELBW individuals were <750 g (27\%), 22\% were <26 weeks, and 24\% were small for gestational age.\(^29\) Mean duration of hospitalization was 101 (SD: 32) days. Both cohorts predominantly were white (>94\%) and from 2-parent families (>79\%), and approximately half were from the upper 2 socioeconomic levels.\(^30\) Highest educational achievement did not differ between cohorts. Mean age at assessment was 23.3 years (SD: 1.2 years) for ELBW and 23.6 years (SD: 1.1 years) for NBW (\(P = .02\)).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ELBW (N = 149)</th>
<th>NBW (N = 133)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth weight, mean (SD), g</td>
<td>841 (124)</td>
<td>3384 (487)</td>
</tr>
<tr>
<td>Gestation, mean (SD), wk</td>
<td>27.1 (2.3)</td>
<td>Term</td>
</tr>
<tr>
<td>&lt;750 g birth weight, n (%)</td>
<td>40 (27)</td>
<td>—</td>
</tr>
<tr>
<td>&lt;26 wk gestation, n (%)</td>
<td>33 (22)</td>
<td>—</td>
</tr>
<tr>
<td>Small for gestational age,(^29) %</td>
<td>36 (24)</td>
<td>4 (3)</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>67 (45)</td>
<td>60 (45)</td>
</tr>
<tr>
<td>White race, n (%)</td>
<td>140 (94)</td>
<td>129 (97)</td>
</tr>
<tr>
<td>Neonatal hospitalization, d</td>
<td>Mean (SD)</td>
<td>101 (32)</td>
</tr>
<tr>
<td>Median (range)</td>
<td>96 (27–193)</td>
<td></td>
</tr>
<tr>
<td>Current family characteristics, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family status (2 parents)</td>
<td>82</td>
<td>79</td>
</tr>
<tr>
<td>Social class(^19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I, II</td>
<td>47</td>
<td>56</td>
</tr>
<tr>
<td>III</td>
<td>30</td>
<td>21</td>
</tr>
<tr>
<td>IV, V</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>YA participants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest educational achievement to date, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>High school</td>
<td>54</td>
<td>56</td>
</tr>
<tr>
<td>College/trades</td>
<td>24</td>
<td>18</td>
</tr>
<tr>
<td>University</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Participants assessed, %</td>
<td>90</td>
<td>92</td>
</tr>
<tr>
<td>Age at assessment, mean (SD), y</td>
<td>23.3 (1.2)</td>
<td>23.6 (1.1)(^*)</td>
</tr>
</tbody>
</table>

\(^*\)\(P = .02\).
Current General Health
There were no significant differences by group among the 8 scales of SF-36, except that ELBW individuals scored lower on mental health \( (P = .04) \) and ELBW male individuals had lower scores than NBW male individuals for physical functioning scale \( (P = .04; \text{Table 2}) \). None of the group \( \times \) gender interactions was significant. There were no differences by group or by gender in the SF-36 summary scores for physical or mental health. However, when asked to rate their health in 1 of 5 categories (excellent, very good, good, fair, or poor), statistically significant differences were found by group \( (P = .03) \) and among male individuals \( (P = .04) \), with ELBW individuals reporting their health less favorably than NBW individuals (and Canadian data\(^{22} \)). These differences became nonsignificant when YAs with NSI were excluded.

Chronic Health Conditions
ELBW YAs had a higher prevalence of NSI (27% vs 2%; \( P < .001^{\text{Holm's}} \)) and were more likely to have multiple (\( \geq 2 \)) impairments (10% vs 0%; \( P = .0004^{\text{Holm's}} \); Table 3). Differences in chronic physical conditions (ELBW versus NBW) were present for the following: seizures (8% vs 2%; OR: 3.8; 95% CI: 1.0–13.7; \( P = .03 \)), fever at any time (6% vs 3%; OR: 2.1; 95% CI: 1.0–4.5; \( P = .03 \)), and recurrent bronchitis (6% vs 1%; OR: 8.5; 95% CI: 1.0–67.9; \( P = .03 \)). Male individuals only; 18% vs 3%; OR: 6.3; 95% CI: 1.3–29.5; \( P = .02 \). The proportion of ELBW individuals with asthma by group and gender is higher than the Canadian data\(^{23} \) (see Table 3, footnote i). There were no differences in emotional problems and mental illness between groups.

When chronic physical health problems were aggregated, ELBW YAs differed significantly from NBW YAs: a lower proportion had no problems, and a higher proportion had \( \geq 3 \) problems \( (P = .01) \). Similar differences were seen in male individuals \( (P < .001^{\text{Holm's}}) \) but not in female individuals \( (P = .89) \). However, when participants with NSI were excluded, these differences no longer were significant for the overall group or for female individuals, but differences among male individuals remained significant \( (P = .001) \). Among those who had at least 1 chronic condition, the mean number per individual was 2.8 (SD: 1.9) for ELBW versus 2.2 (SD: 0.14) for NBW \( (P = .01) \).

Injuries
Self-report of injuries in the past 12 months, serious enough to limit normal activities, did not differ between groups (ELBW 24% vs NBW 22%; \( P = .71 \)). However, both groups reported slightly higher rates than the general Canadian population (18%, aged 20–24).\(^{23} \)

Late Retinal Detachment
In eliciting details of types of surgeries, we unexpectedly found that 6 ELBW YAs (4%) experienced sudden late retinal detachment, with 2 YAs remaining blind in the affected eye after laser surgery. Voluntary ophthalmological examinations became nonsignificant when YAs with NSI were excluded.

| TABLE 2 | SF-36 Scores: 8 Scales and Summary Scores for ELBW and NBW YAs According to Gender |
| --- | --- | --- | --- | --- | --- |
| Parameter | ELBW \( (N = 62) \), Mean (SD)\(^a \) | NBW \( (N = 60) \), Mean (SD) | P | ELBW \( (N = 80) \), Mean (SD) | NBW \( (N = 73) \), Mean (SD) | P |
| Physical functioning | 54.9 (5.1) | 56.3 (1.6) | .04 | 53.3 (6.1) | 54.1 (5.4) | .39 | 54.0 (5.7) | 55.1 (4.2) | .07 |
| Role—physical | 53.3 (7.9) | 54.1 (6.0) | .50 | 53.2 (7.1) | 52.0 (8.3) | .32 | 53.2 (7.4) | 53.0 (7.4) | .74 |
| Bodily pain | 55.1 (8.1) | 55.4 (8.1) | .84 | 53.8 (9.6) | 54.3 (9.5) | .73 | 54.4 (9.0) | 54.8 (8.9) | .68 |
| General health | 54.0 (10.0) | 56.3 (6.7) | .12 | 51.3 (11.4) | 52.3 (10.4) | .54 | 52.4 (10.8) | 54.1 (9.1) | .16 |
| Vitality | 54.4 (10.5) | 55.6 (7.7) | .48 | 51.4 (10.8) | 53.6 (9.7) | .18 | 52.7 (10.7) | 54.5 (8.9) | .13 |
| Social functioning | 51.1 (9.4) | 51.4 (7.8) | .85 | 49.6 (11.2) | 51.1 (7.8) | .32 | 50.2 (10.5) | 51.2 (7.7) | .36 |
| Role—emotional | 50.0 (10.6) | 51.8 (5.8) | .24 | 49.3 (10.8) | 50.8 (9.3) | .37 | 49.6 (10.7) | 51.2 (7.9) | .15 |
| Mental health | 50.2 (12.0) | 52.6 (8.7) | .20 | 48.5 (12.9) | 51.4 (8.9) | .10 | 49.2 (12.5) | 51.9 (8.8) | .04 |
| Summary measures | | | | | | | | | |
| Physical health\(^b \) | 55.9 (6.6) | 56.7 (5.2) | .50 | 54.5 (7.0) | 54.0 (7.9) | .68 | 55.1 (6.9) | 55.2 (6.9) | .92 |
| Mental health\(^c \) | 49.4 (12.0) | 51.2 (8.7) | .35 | 47.7 (13.6) | 50.7 (9.8) | .12 | 48.4 (12.9) | 50.9 (9.3) | .07 |
| Self-rated health, n (%) | Canada\(^a \), % | | Canada\(^a \), % | | Canada\(^a \), % | | Canada\(^a \), % | | |
| Excellent | 20 (32) | 19 (32) | 27 | 17 (21) | 15 (21) | 27 | 37 (26) | 34 (26) | 27 |
| Very good | 16 (26) | 29 (48)\(^d \) | 42 | 34 (43) | 35 (48) | 43 | 50 (35) | 64 (48)\(^d \) | 43 |
| Good | 18 (29) | 9 (15)\(^d \) | 25 | 20 (25) | 19 (26) | 24 | 38 (27) | 28 (21)\(^d \) | 25 |
| Fair/poor | 8 (13) | 3 (5)\(^d \) | 5 | 9 (11) | 4 (5) | 6 | 17 (12) | 7 (5)\(^d \) | 5 |

\(^a\) Excludes 7 ELBW YAs with severe disabilities, who were unable to answer for themselves (male, \( n = 5 \); female, \( n = 2 \)).

\(^b\) Norm-based: mean \( = 50 \), SD \( = 10 \), based on 1998 US general population data.\(^{26} \)

\(^c\) Derived from physical functioning, role—physical (limitations as a result of physical health), bodily pain, and general health.

\(^d\) Derived from vitality, social functioning, role—emotional (limitations as a result of emotional problems), and mental health.

\(^e\) Canadian data from Statistics Canada\(^{22} \) (ages 20–24 years, \( n = 6 \) 404 036).

\(^f\) Comparison of self-rated health ELBW versus NBW: male \( P = .04 \); female \( P = .35 \); group \( P = .03 \); when individuals with NSI were excluded, all \( P \) values were nonsignificant.

\(^g\) Data on categories fair and poor from the self-reported health status were combined because of small numbers and for comparison with similarly classified Canadian data.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Male ELBW (N = 67), n (%)</th>
<th>Female ELBW (N = 82), n (%)</th>
<th>P</th>
<th>Male NBW (N = 60), n (%)</th>
<th>Female NBW (N = 73), n (%)</th>
<th>P</th>
<th>Male Total (N = 149), n (%)</th>
<th>Female Total (N = 133), n (%)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSI</td>
<td>18 (27)</td>
<td>0 (0)</td>
<td>&lt;.001c</td>
<td>22 (27)</td>
<td>3 (4)</td>
<td>&lt;.001c</td>
<td>40 (27)</td>
<td>3 (2)</td>
<td>&lt;.001c</td>
</tr>
<tr>
<td>Cerebral palsy</td>
<td>11 (16)</td>
<td>0 (0)</td>
<td>.001c</td>
<td>9 (11)</td>
<td>1 (1)</td>
<td>.01</td>
<td>20 (13)</td>
<td>1 (1)</td>
<td>.001c</td>
</tr>
<tr>
<td>Hydrocephalus</td>
<td>3 (4)</td>
<td>.24</td>
<td></td>
<td>3 (4)</td>
<td>0 (0)</td>
<td>.24</td>
<td>6 (4)</td>
<td>0 (0)</td>
<td>.03</td>
</tr>
<tr>
<td>Microcephaly</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>1.0</td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>1.0</td>
</tr>
<tr>
<td>Mental retarditation×</td>
<td>6 (10)</td>
<td>0 (0)</td>
<td>.02</td>
<td>6 (7)</td>
<td>1 (1)</td>
<td>.12</td>
<td>12 (8)</td>
<td>1 (1)</td>
<td>.004d</td>
</tr>
<tr>
<td>Autism</td>
<td>2 (3)</td>
<td>0 (0)</td>
<td>.49</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td>2 (1)</td>
<td>0 (0)</td>
<td>50</td>
</tr>
<tr>
<td>Blindness</td>
<td>5 (7)</td>
<td>0 (0)</td>
<td>.03</td>
<td>14 (17)</td>
<td>0 (0)</td>
<td>.0002c</td>
<td>19 (13)</td>
<td>0 (0)</td>
<td>.0006ec</td>
</tr>
<tr>
<td>Deafness</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td>0 (0)</td>
<td>1 (1)</td>
<td>.95</td>
<td>0 (0)</td>
<td>1 (1)</td>
<td>95</td>
</tr>
<tr>
<td>Multiple impairments (≥2)</td>
<td>7 (10)</td>
<td>0 (0)</td>
<td>.01</td>
<td>8 (10)</td>
<td>0 (0)</td>
<td>.006</td>
<td>15 (10)</td>
<td>0 (0)</td>
<td>.0004ec</td>
</tr>
</tbody>
</table>

**Physical health**

- Seizures: 3 (4) vs. 3 (4), 1.7
- Allergies: 28 (34) vs. 31 (42), .28
- Asthma: 18 (22) vs. 15 (21), .83
- Cardiac problems: 3 (4) vs. 2 (3), 1.0
- Recurrent bronchitis: 4 (5) vs. 1 (1), .37
- Diabetes: 0 (0) vs. 1 (1), .47
- Migraines: 20 (24) vs. 14 (19), .43
- Back problems: 12 (15) vs. 20 (27), .05
- Other problems: 18 (22) vs. 13 (18), .52
- Mental health:
  - Emotional problems: 18 (22) vs. 14 (19), .67
  - Mental illness: 8 (10) vs. 11 (11), .80
- Chronic health conditions, total:
  - None: 16 (24) vs. 32 (53), 21 (27) vs. 17 (23)
  - 1–2: 29 (43) vs. 26 (43), .001c
  - ≥3: 22 (33) vs. 2 (5)
- Chronic conditions ≥ 1/YA:
  - Mean (SD): 2.5 (1.6) vs. 1.6 (1.7)
  - Range: 1–7 vs. 1–4

**Notes:**
- Includes parental proxy responses for 7 ELBW YAs with severe impairment, who were unable to respond.
- Diagnosed by a health professional, not mutually exclusive; some individuals had ≥ 1 impairment.
- Significant after Holm’s correction.
- OR (95% CI) ELBW versus NBW: cerebral palsy, 20.5 (2.7–154.7); mental retardation, 11.6 (1.4–90.1); seizures, 3.8 (1.0–13.7); asthma, male individuals, 6.3 (1.3–29.5); recurrent bronchitis, 8.5 (1.0–67.9).
- Defined as Wechsler Intelligence Scale for Children–Revised IQ < 70 at age 8 years.
- Eight unilaterally blind and 11 bilaterally blind ELBW individuals.
- One unilaterally deaf NBW individual.
- Self-reported health conditions, diagnosed by a health professional and lasting longer than 6 months; not mutually exclusive; some individuals had ≥ 1 problem. Food allergies excluded from “allergies” variable.
- Includes lupus, multiple sclerosis, scoliosis, hepatitis C, gastrointestinal problems, and arthritis; total number includes these and other health conditions that were not listed because of low incidence.
logic assessments on 45 ELBW YAs yielded 3 additional cases of retinal tears that also required surgery. None of the NBW YAs experienced the same.

Current Functional Limitations
ELBW YAs reported significantly more functional limitations than NBW YAs by group and by gender for the following variables (Table 4): difficulty seeing (group P < .001Holm’s, male P = .002Holm’s, female P < .001Holm’s), bilateral blindness (group P = .001Holm’s, female P = .01), clumsiness (group P = .001Holm’s; male P = .007, female P = .02), dexterity (P = .002Holm’s; male P = .02, female P = .02), and learning disabilities (group P < .001Holm’s, male P = .003, female P < .001Holm’s). In addition, there were differences by group (P = .04) but not by gender for hearing difficulties and reduced self-care abilities (P = .03). These results include 7 parental proxy responses.

When functional limitations were summed and compared by categories of none, 1 to 2 and ≥3, ELBW individuals were less likely to have no limitations and more likely to have multiple limitations. These differences were significant by group (P < .001Holm’s) and by gender (P < .001Holm’s) and remained significant when NSI were excluded.

More ELBW than NBW YAs reported limitations in carrying out “normal daily activities” as a result of health problems (21% vs 11%; OR: 2.1; 95% CI: 1.0–4.0; P = 0.03, NS by gender; corresponding Canadian data 19.1%). These differences became NS when individuals with NSI were excluded. Among those with limitations, there were significant differences between groups in the reasons for limitations (P = .00004, data not shown): a significant majority (81%) of ELBW individuals cited mental illness and NSI as the main reasons, whereas a similar proportion (80%) of NBW individuals identified chronic conditions and acute injuries.

There were no significant differences by group (ELBW versus NBW: 24% vs 19%) or by gender (male: 16% vs 13%; female: 30% vs 23%) in the proportion with any absenteeism from school/work as a result of illness during the previous month. The proportions with absenteeism are similar to the Canadian data (male: 14.7%; female: 22.2%; total: 18.3%). However, significant differences were noted for mean number of days absent among female individuals (ELBW: 5.5 days [SD: 6.3 days]; NBW: 2.6 days [SD: 2.2 days]; P = .04, data not shown).

Health Care Use
There were no significant differences by group or by gender for the proportion with overnight hospitalizations or surgery in the past 12 months (Table 5). During the past 6 months, there were no differences in the proportion with visits to the emergency department or visits to any health professionals by group or by gender, with the exception of social worker contacts by ELBW female individuals (13% vs 4%; P = .04). The proportion who had outpatient investigations did not differ by group or by gender.

The use of prescription drugs differed significantly by group only for medications for depression (ELBW 14% vs NBW 6%; OR: 2.6; 95% CI: 1.0–6.0; P = .025) but not by gender. In terms of assistive devices/aids, there were no significant differences in the use of braces, crutches, or canes between groups, although the reasons for their use differed: ELBW individuals had chronic conditions related to NSI, whereas NBW individuals had acute injuries. Four ELBW and 1 NBW YA required wheelchairs, whereas none of the NBW individuals did (P = .43); and prescription glasses were required more often by ELBW YAs (64% vs 37%; OR: 3.1; 95% CI: 1.9–5.0; P < .001Holm’s; male P = .002, female P < .001Holm’s). In addition, a higher proportion of ELBW individuals required visual aids such as Braille equipment and canes (group P = .002Holm’s; female P = .003Holm’s). Very few participants wore hearing aids (2 female ELBW and 1 female NBW), despite more who reported “hearing problems.”

Although a minority of individuals required home-care services, there were significant differences between groups (8% vs 2%; OR: 5.7; 95% CI: 1.2–26.1; P = .01). Significantly more ELBW than NBW YAs received services such as household help, specialized companionship, and personal care (7% vs 1%; OR: 9.5; 95% CI: 2.0–39.5; P = .009). One ELBW female individual with severe impairment required episodic respite care, and another ELBW female individual was in permanent foster care.

Physical Self-Efficacy, Physical Activity, and Hand-Grip Strength
ELBW YAs had significantly lower total scores in the physical self-efficacy scale (P < .001Holm’s) and in the 2 subscales of perceived physical ability (P < .001Holm’s) by group and by gender and in physical self-presentation confidence by group (P = .001Holm’s) and by female gender (P = .002Holm’s; Table 6). These differences persisted even when individuals with NSI were excluded.

A significantly lower proportion of ELBW versus NBW YAs reported regular participation in sports and strenuous activities (38% vs 59%; P = .001Holm’s), and the proportion also was lower in comparison with Canadian data (60.2%; see Table 6, footnote d). A higher proportion of ELBW individuals attributed the lower participation rates to health conditions (22% vs 9%; P = .004Holm’s). However, although differences were observed among male individuals (P < .001Holm’s), there were no significant differences in the proportion of female individuals who participated and those who were unable to participate as a result of health conditions. The
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ELBW</strong> (N = 67), n (%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty seeing</td>
<td>36 (54)</td>
<td>16 (27)</td>
<td>&lt;0.002</td>
</tr>
<tr>
<td>Blindness, bilateral</td>
<td>3 (4)</td>
<td>0 (0)</td>
<td>0.28</td>
</tr>
<tr>
<td>Difficulty hearing&lt;sup&gt;d&lt;/sup&gt;</td>
<td>3 (4)</td>
<td>1 (2)</td>
<td>0.62</td>
</tr>
<tr>
<td>Emotional problems</td>
<td>8 (12)</td>
<td>4 (7)</td>
<td>0.31</td>
</tr>
<tr>
<td>Mental illness</td>
<td>4 (6)</td>
<td>1 (2)</td>
<td>0.36</td>
</tr>
<tr>
<td>Mobility (unable to walk unaided)</td>
<td>2 (3)</td>
<td>0 (0)</td>
<td>0.49</td>
</tr>
<tr>
<td>Clumsiness</td>
<td>8 (12)</td>
<td>0 (0)</td>
<td>0.007</td>
</tr>
<tr>
<td>Dexterity</td>
<td>6 (9)</td>
<td>0 (0)</td>
<td>0.02</td>
</tr>
<tr>
<td>Learning disabilities</td>
<td>18 (27)</td>
<td>4 (7)</td>
<td>0.03</td>
</tr>
<tr>
<td>Hyperactivity</td>
<td>6 (9)</td>
<td>2 (3)</td>
<td>0.27</td>
</tr>
<tr>
<td>Pain/discomfort</td>
<td>14 (21)</td>
<td>7 (12)</td>
<td>0.16</td>
</tr>
<tr>
<td>Reduced self-care abilities</td>
<td>4 (6)</td>
<td>0 (0)</td>
<td>0.12</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limitation in daily activities&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At home/school/work as a result of health problems</td>
<td>12 (18)</td>
<td>4 (7)</td>
<td>0.05</td>
</tr>
</tbody>
</table>

<sup>a</sup> Includes parent proxy respondents for 7 YAs with severe disabilities, who were unable to answer for themselves.

<sup>b</sup> All wear glasses; includes 8 ELBW individuals who are unilaterally blind.

<sup>c</sup> Significant after Holm’s correction.<sup>28</sup>

<sup>d</sup> Includes 1 female NBW individual who had a diagnosis of unilateral deafness and 2 female ELBW individuals who had hearing loss that required hearing aids.

<sup>e</sup> Sum of 12 problems listed above.

<sup>f</sup> When NSI were excluded, the differences remained significant by group and by gender: group <0.001, male <0.001, and females <0.006.

<sup>g</sup> Canadian data from Statistics Canada<sup>23</sup> (aged 20–24, self-report). Participation and activity limitations: male 18%; female 21%; total 19%.

<sup>h</sup> Not significant when NSI excluded; P = 0.61.
TABLE 5  Health Care Use in Past 6 to 12 Months: ELBW and NBW YAs According to Gender

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Male ELBW (N = 67, n %)</th>
<th>Female ELBW (N = 82, n %)</th>
<th>P</th>
<th>Male NBW (N = 60, n %)</th>
<th>Female NBW (N = 73, n %)</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past 12 mo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitalization overnight</td>
<td>4 (6)</td>
<td>0 (0)</td>
<td>.12</td>
<td>11 (13)</td>
<td>12 (16)</td>
<td>.59</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>No. of times, mean (SD)</td>
<td>1.5 (1.0)</td>
<td>0 (0)</td>
<td>.10</td>
<td>1.1 (3.0)</td>
<td>1.5 (7.9)</td>
<td>.12</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>No. of nights, mean (SD)</td>
<td>5.8 (9.5)</td>
<td>0 (0)</td>
<td>.49</td>
<td>6.5 (6.5)</td>
<td>4.5 (3.9)</td>
<td>.38</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Surgery</td>
<td>6 (9)</td>
<td>3 (5)</td>
<td></td>
<td>8 (10)</td>
<td>6 (8)</td>
<td>.73</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Past 6 mo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency department</td>
<td>15 (22)</td>
<td>8 (13)</td>
<td>.17</td>
<td>18 (22)</td>
<td>12 (16)</td>
<td>.40</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Family physician</td>
<td>37 (55)</td>
<td>26 (43)</td>
<td>.18</td>
<td>62 (76)</td>
<td>56 (77)</td>
<td>.87</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Specialist physician</td>
<td>16 (24)</td>
<td>9 (15)</td>
<td>.19</td>
<td>23 (28)</td>
<td>23 (32)</td>
<td>.63</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Psychologist</td>
<td>3 (4)</td>
<td>4 (7)</td>
<td>.70</td>
<td>4 (5)</td>
<td>2 (3)</td>
<td>.68</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Social worker</td>
<td>1 (1)</td>
<td>2 (3)</td>
<td>.60</td>
<td>11 (13)</td>
<td>3 (4)</td>
<td>.04</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Physiotherapist/OT</td>
<td>3 (4)</td>
<td>2 (3)</td>
<td>.10</td>
<td>5 (6)</td>
<td>7 (10)</td>
<td>.41</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Speech therapist</td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>1.0</td>
<td>0 (0)</td>
<td>1 (1)</td>
<td>.47</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Other health care providers</td>
<td>7 (10)</td>
<td>5 (8)</td>
<td>.68</td>
<td>10 (12)</td>
<td>14 (19)</td>
<td>.23</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Outpatient tests</td>
<td>18 (27)</td>
<td>14 (23)</td>
<td>.64</td>
<td>38 (46)</td>
<td>32 (44)</td>
<td>.65</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Prescription medications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(currently required on a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>regular basis)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epilepsy</td>
<td>3 (4)</td>
<td>0 (0)</td>
<td>.24</td>
<td>5 (6)</td>
<td>3 (4)</td>
<td>.72</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Hyperactivity/behavioral</td>
<td>2 (3)</td>
<td>0 (0)</td>
<td>.52</td>
<td>11 (13)</td>
<td>3 (4)</td>
<td>.04</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>other</td>
<td>5 (7)</td>
<td>1 (2)</td>
<td>.21</td>
<td>16 (20)</td>
<td>7 (10)</td>
<td>.41</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Depression</td>
<td>5 (7)</td>
<td>3 (5)</td>
<td>.72</td>
<td>8 (10)</td>
<td>5 (7)</td>
<td>.51</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Asthma</td>
<td>5 (7)</td>
<td>3 (5)</td>
<td>.72</td>
<td>8 (10)</td>
<td>5 (7)</td>
<td>.51</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Assistive devices/aids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Required in past 12 mo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(not mutually exclusive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Braces/criutches/cane</td>
<td>3 (4)</td>
<td>4 (7)</td>
<td>.88</td>
<td>4 (5)</td>
<td>3 (2)</td>
<td>.72</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Wheelchair</td>
<td>3 (4)</td>
<td>0 (0)</td>
<td>.28</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>1.0</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Prescription glasses/contacts</td>
<td>36 (54)</td>
<td>16 (27)</td>
<td>.02</td>
<td>60 (73)</td>
<td>33 (45)</td>
<td>.003</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Vision aids (Braille</td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>1.0</td>
<td>9 (11)</td>
<td>0 (0)</td>
<td>1.0</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>equipment, white cane, etc.</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1.0</td>
<td>3 (4)</td>
<td>1 (1)</td>
<td>.62</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Hearing aids</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1.0</td>
<td>3 (4)</td>
<td>1 (1)</td>
<td>.62</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Other adaptations (house</td>
<td>2 (3)</td>
<td>0 (0)</td>
<td>.52</td>
<td>4 (5)</td>
<td>1 (1)</td>
<td>.43</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>/furniture/car)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home care (past 6 mo)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Required any home care</td>
<td>4 (6)</td>
<td>0 (0)</td>
<td>.12</td>
<td>8 (10)</td>
<td>2 (2)</td>
<td>.10</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>service</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of services required</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(not mutually exclusive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skilled medical care</td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>1.0</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>1.0</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Personal care</td>
<td>2 (3)</td>
<td>0 (0)</td>
<td>.49</td>
<td>2 (2)</td>
<td>0 (0)</td>
<td>.49</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Other (companionship,</td>
<td>5 (7)</td>
<td>0 (0)</td>
<td>.08</td>
<td>5 (6)</td>
<td>1 (1)</td>
<td>.26</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>household help)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporary respite care</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>1.0</td>
<td>.001Holm's</td>
</tr>
<tr>
<td>Permanent foster care</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>1.0</td>
<td>.001Holm's</td>
</tr>
</tbody>
</table>

OT indicates occupational therapist.

a Includes parental proxy responses for 7 ELBW individuals with severe disabilities, who were unable to answer for themselves.

b When child birth was excluded from those with overnight hospitalization, the proportions decreased to 10% for female individuals in both groups.

c OR (95% CI): Social worker, female 3.6 (0.9 – 13.5); any home care service, total group 5.7 (1.2 – 26.1); other home care services, total group 9.0 (2.0 – 39.5); medications for depression, total group 2.6 (1.0 – 6.0); prescription glasses, male 3.2 (1.5 – 6.7), females 3.3 (1.6 – 6.4), total group 3.1 (1.9 – 5.0).

d For example, chiropractor, massage therapist, optometrist.

e Some mobility aids were required for temporary only (eg, orthopedic surgery). One ELBW female with hearing problems required a telephone device, this is included in the total with hearing aids. An additional 6 ELBW required minor aids (not shown in table), eg, orthotics, glucose monitor, ostomy supplies, special utensils, etc.

f Significant after Holm's correction.\(^{28}\)

differences between groups in participation rates remained significant after NSI were excluded (P = .02).

Despite exclusion of those with NSI, ELBW YAs had significantly lower hand-grip strength in their dominant hand compared with NBW YAs (32 [SD: 10] vs 38 [SD: 10]; P < .001Holm's) and by gender (male P = .002Holm's; female P < .001Holm's). The analysis of variance was significant for group (P < .001) and gender (P < .001), but there was no interaction.

**DISCUSSION**

The findings of this study confirm our hypotheses that by the time ELBW survivors reach adulthood, their current physical and mental health is similar to that of NBW YAs for overall group and by gender. Except for a higher prevalence of NSI, differences were observed in only a few residual chronic health conditions, such as seizures, recurrent bronchitis, and asthma (male individuals). Although still significantly different, both groups had a
Physical Self-Efficacy, Participation in Physical Activity, and Hand-Grip Strength by Gender Among ELBW and NBW YAs

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
<th>Mean Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ELBW (N = 62)</td>
<td>NBW (N = 60)</td>
<td>P</td>
<td>ELBW (N = 80)</td>
</tr>
<tr>
<td>Physical Self-Efficacy Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score, mean (SD)</td>
<td>85 (14)</td>
<td>93 (14)</td>
<td>&lt;001b</td>
<td>88 (16)</td>
</tr>
<tr>
<td>Subscales, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived physical ability</td>
<td>41 (9)</td>
<td>46 (8)</td>
<td>0.01b</td>
<td>37 (8)</td>
</tr>
<tr>
<td>Physical self-presentation confidence</td>
<td>51 (10)</td>
<td>54 (8)</td>
<td>.13</td>
<td>48 (8)</td>
</tr>
<tr>
<td>Participation in physical activities</td>
<td>67 (110)</td>
<td>60 (100)</td>
<td>.001b</td>
<td>82 (110)</td>
</tr>
<tr>
<td>Regular participation in sports/tremendous activities, n (%)</td>
<td>31 (46)</td>
<td>46 (77)</td>
<td>&lt;0.01b</td>
<td>26 (32)</td>
</tr>
<tr>
<td>Unable to participate as a result of health conditions, n (%)</td>
<td>15 (22)</td>
<td>1 (2)</td>
<td>&lt;0.01b</td>
<td>17 (21)</td>
</tr>
<tr>
<td>Hand-grip strength</td>
<td>41 (9)</td>
<td>47 (9)</td>
<td>&lt;002b</td>
<td>25 (5)</td>
</tr>
</tbody>
</table>

Univariate ANOVA by group and gender for physical self-efficacy scale and hand grip. Total score: group P < 0.01, gender P < 0.01, interaction P = .89; perceived physical ability: group P < 0.01, gender P < 0.01, interaction P = .43; physical self-presentation confidence: group P = .001, gender P = .02, interaction P = .33; hand grip: group P < 0.01, gender P < 0.01, interaction P = .87.

a Excludes 7 ELBW YAs with severe disabilities, who were unable to respond for themselves.

b Significant after Holm’s correction.

c Includes parental proxy responses for 7 ELBW YAs with severe disabilities.

d Canadian data from Statistics Canada23 (aged 20–24 years). Physically active or moderately physically active: male 64%; female 56%; total 60%.

e Dominant hand of YA without NSI.
believe that there may be other undetected cases. Similarly, the 15-year outcome study of the Cryotherapy for Retinopathy of Prematurity trial for threshold retinopathy found retinal detachment in 4.5% of treated eyes and 7.7% of control eyes. What is worrisome is that these events occurred in eyes that were judged to be normal at the 10-year assessment. The development of these adverse outcomes indicates the need for lifelong follow-up of people with a history of retinopathy.

There are no available reports in the literature for comparison with our study regarding use of health care resources by ELBW individuals at young adulthood. Direct comparison with our same cohort at adolescence was not possible because the respondents in the last study were parents and the time frame was in the past 2 years. The respondents this time were the YAs, and for reasons of recall bias, the period of inquiry was only for the last 6 to 12 months. We are aware of differences in perceptions regarding health conditions between parents and children, but whether there also are differences in responses related to health care use is not known. It is possible that the lack of significant differences in this study may be a reflection of the overall low base rates of health care use at this age. It is clear that the costs to the health care system and other services that are associated with extreme prematurity are substantial in the early years and persist into midchildhood. Thereafter, we have shown that there is a significant decline in use of health services around adolescence and beyond. Contrary to the earlier pessimistic projections, ELBW YAs do not pose a considerable lifelong burden to the health care system. Economic evaluation is in progress.

Consistent with this study, several investigators reported that VLBW adolescents and YAs lead a less active physical lifestyle and have limited participation in sports and strenuous activities. A reduction in muscle strength and physical working capacity among VLBW boys was reported by Ericson et al. Rogers et al found that 17-year-old unimpaired ELBW survivors had significantly lower motor performance than control subjects in aerobic capacity, strength, endurance, flexibility, and activity level. They speculated whether the lower scores on these measures were a result of extreme prematurity (subclinical pulmonary compromise or subtle neuromotor difficulties) or possible sheltering by parents or reflected a preference by the ELBW individuals for a lower physically active lifestyle. However, we and others have reported that ELBW children are consistently described by their parents as having problems with clumsiness and coordination, which also were acknowledged by the ELBW YAs themselves in this study. In addition, they rated themselves lower in perceived physical ability and physical self-presentation confidence and were found to have lower hand-grip strength. Overall, a lower proportion of ELBW YAs participated in regular physical activities in comparison with NBW YAs and the Canadian national norms.

CONCLUSIONS
This is the first longitudinal study of the health status, chronic conditions, physical abilities, and health care use at young adulthood for ELBW survivors who were born in the early era of neonatal intensive care. These data were derived from a defined geographic region with high participation rates, include a control group, and were collected by experienced interviewers who were masked to the group status. We acknowledge that there may be recall bias inherent in self-reported data among both groups, which were not corroborated by physician or hospital charts. Furthermore, although this is the largest study to date on health outcomes at young adulthood, the number of subjects in epidemiologic terms is relatively small. It is possible that we may have been underpowered to detect differences in some outcomes. Nevertheless, the findings for ELBW YAs are optimistic in terms of a reasonably good current health status and a significant decline in use of health care resources. However, as anticipated, ELBW YAs continue to suffer from chronic health conditions, and a significantly higher proportion have functional limitations. For various reasons, ELBW YAs seem to lead a more sedentary lifestyle than their peers. Additional follow-up to late adulthood is essential to determine whether the chronic health conditions and functional limitations among ELBW YAs will get progressively worse with age and whether they will have a higher prevalence of cardiovascular disease and metabolic problems in the future. Finally, although these data may not be entirely generalizable to the current, even smaller survivors of neonatal intensive care, they provide some guidelines to health care providers and funding agencies for projection, planning, and allocation of the necessary resources for future at-risk infants.

ACKNOWLEDGMENTS
This study was supported by grants MOP42536 from the Canadian Institutes of Health Research and 1 RO1 HD40219 from the National Institute of Child Health and Human Development.

We thank the ELBW and NBW YAs and their parents for cooperation with our many studies. We also thank our research staff Liz Merz (for tracing the participants), Lorraine Hoult and Mary Lou Schmuck (for statistical analysis), and Diane Turcotte (for typing the manuscript). We appreciate the support of the department of Pediatrics and the Children’s Hospital, McMaster University (Hamilton, ON, Canada).

REFERENCES


19. Hollingshead AS. Four Factor Index of Social Status. New Haven, CT: Yale University; 1975


35. Petrou S. The economic consequences of preterm birth during the first 10 years of life. BJOG. 2005;112:10–15


37. Rogers M, Fay TB, Whitfield MF, Tomlinson J, Grunau RE. Aerobic capacity, strength, flexibility, and activity level in impaired extremely low birth weight (<800 g) survivors at 17 years of age compared with term-born control subjects. Pediatrics. 2005;116(1). Available at: www.pediatrics.org/cgi/content/full/116/1/e58


40. Barker DJ, Bull AR, Simmonds SJ. Fetal and placental size and risk of hypertension in adult life. BMJ. 1990;301:259–262

41. Irving RJ, Belon NR, Elton RA, Walter BR. Adult cardiovas-


Measles-Mumps-Rubella and Varicella Vaccine Responses in Extremely Preterm Infants

Carl T. D’Angio, MD, Paulina A. Boohene, MD, Anne Mowerer, RN, Susette Audet, BS, Marilyn A. Menegus, PhD, D. Scott Schmid, PhD, Judy A. Beeler, MD

Departments of Medicine, Pediatrics, and Microbiology and Immunology, University of Rochester, Rochester, New York; Center for Biologics Evaluation and Research, US Food and Drug Administration, Bethesda, Maryland; National VZV Laboratory, Centers for Disease Control and Prevention, Atlanta, Georgia

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

ABSTRACT

OBJECTIVE. Extremely preterm infants mount lower antibody responses than term infants to several vaccines. The objective of this study was to measure the immunogenicity of measles-mumps-rubella and varicella vaccines in preterm and term children.

METHODS. Immune status before immunization and immune response after immunization with measles-mumps-rubella and varicella vaccines at 15 months of age were compared in 32 infants, 16 of whom were preterm (<29 weeks’ gestation) and 16 of whom were term (≥37 weeks’ gestation) at birth. Blood was drawn before vaccination and 3 to 6 weeks thereafter. Measles antibody was measured by plaque reduction neutralization assay. Mumps and rubella immunoglobulin G were measured in available sera by enzyme-linked fluorescent immunoassay. Varicella immunoglobulin G was measured in available sera by glycoprotein enzyme-linked immunosorbent assay. Values that were above or below the assay limits were assigned values double or half those limits, respectively. The primary outcome was the geometric mean antibody titer.

RESULTS. Preterm children had lower mumps and rubella geometric mean titers than did term children before vaccine, and nearly all children were seronegative for each of the 4 vaccine antigens before immunization. Measles, mumps, rubella, and varicella geometric mean titers were similar between groups after vaccine. All children were seropositive for measles after vaccine, whereas 13 of 14 preterm and 11 of 13 term children were seropositive for mumps, 13 of 14 preterm and 13 of 13 term children were seropositive for rubella, and 11 of 16 preterm and 9 of 15 term children were seropositive for varicella.

CONCLUSIONS. Preterm children mounted antibody responses that were similar to those of term children after measles-mumps-rubella and varicella vaccines at 15 months of age.
THE AMERICAN ACADEMY OF PEDIATRICS recommends that preterm infants receive most vaccines, including the measles-mumps-rubella (MMR) and varicella vaccines, at the same chronological age as term infants.\(^1\) However, extremely preterm infants (born at \(<28–30\) weeks’ gestation) have lower antibody responses than term infants to several vaccines, including \textit{Haemophilus influenzae} type b,\(^2\) hepatitis B,\(^3\) and polio,\(^12,\) when these vaccines are given at the postnatal ages that are recommended for term infants. The immunogenicity of vaccine series that are begun during the first 6 months in preterm infants may remain diminished after the administration of toddler\(^14,\) and school-age\(^15,\) boosters. Inactivated influenza vaccine, which is administered at or after 6 months of age, also seems to be less immunogenic in preterm than term infants.\(^17\) These data raise the question of whether a relative deficit in humoral immunity persists beyond 6 months’ postnatal age in preterm infants.

We evaluated the serologic response of extremely preterm infants to MMR and varicella vaccines that were given at 15 months’ chronological age. We hypothesized that these children would have lower geometric mean antibody titers (GMT) to varicella, mumps, measles, and rubella after vaccination than infants who were born at term but that a similar number of preterm as term infants would become seropositive.

METHODS

Patients

The study was conducted between May 2002 and May 2005. The University of Rochester Institutional Review Board approved the study. Infants were eligible for the study when they were born at \(<29\) weeks’ or \(\geq37\) weeks’ gestation, they were \(<16\) months of age, they had not yet received MMR or varicella vaccines, and the parents had given permission for and the primary pediatrician had agreed to the study. Infants were excluded when they had known immunodeficiency, had contraindications to vaccination, or required systemic corticosteroids or oxygen therapy at the time of vaccination. Concurrent administration of other vaccines was permitted.

Preterm infants were recruited at the time of a scheduled visit to the regional perinatal center’s neurodevelopmental follow-up clinic (at which all infants who were born at \(<30\) weeks’ gestation were evaluated) at ~9 to 12 months of age. Term infants were recruited from either a hospital-based pediatric clinic or a private pediatric practice at the time of the 12-month visit. Term infants were matched to preterm infants on race and, whenever possible, on gender.

Study Visits

Written parental permission, demographic and historical information, and a blood specimen were obtained before immunization. MMR II (Merck Vaccines, Whitehouse Station, NJ) and varicella (Varivax; Merck) vaccines were administered by the primary pediatrician from his or her routine office stock according to the manufacturers’ instructions and routine medical care practices. At a visit 3 to 6 weeks after immunization, interim medical history and a blood specimen were obtained.

Serology

Serum was separated and frozen, in 3 aliquots, at \(-80^\circ\)C until the time of analysis. Measles antibody was measured by plaque reduction neutralization assay according to established protocols.\(^18,\) Briefly, measles virus (25–30 plaques per well) and serial dilutions of heat-inactivated subject serum were incubated at 36°C, added to Vero cell monolayers in 24-well culture plates in duplicate, and covered with carboxymethylcellulose overlay media. The trays then were incubated for 5 days at 36°C, at which time the monolayers were stained and fixed and the plaques were counted. The end point for the test was calculated using the Kärber formula and reflected the dilution of serum that reduced the number of plaques by 50%. Dilution values were converted to mIU by comparison with results that were obtained using second International Standard (66/202) tested in parallel; in this assay, a titer of 1:8 was equivalent to 8 mIU/mL. A neutralizing antibody titer \(\geq120\) mIU/mL was considered evidence of seropositivity.\(^20\)

Mumps and rubella immunoglobulin G (IgG) titers were measured using a commercially available, automated, enzyme-linked fluorescent immunoassay (VIDAS Vitek ImmunoDiagnostic Assay System; bioMérieux, Inc, Hazelwood, MO). Titers were reported as relative fluorescence values (RFV). Mumps values \(<0.35\) were considered negative, values 0.35 to 0.49 were considered equivocal (and treated as negative during analysis), and values \(\geq0.50\) were considered positive. Rubella values \(<0.40\) were considered negative, values 0.40 to 0.49 were considered equivocal (and treated as negative during analysis), and values \(\geq0.50\) were considered positive.

Varicella IgG was measured using a glycoprotein enzyme-linked immunosorbent assay (ELISA) protocol that essentially was the same as that reported by Wasmuth and Miller.\(^21\) A twofold end point dilution series that began with a dilution of 1:20 was performed. Results for each dilution were scored as positive, negative, or equivocal on the basis of empirically determined cutoffs that were based on calculations of cumulative variance of test data from repeated testing of known varicella zoster virus–positive and –negative sera by several operators. Titration end points were defined as the highest dilution to produce a test result in the equivocal range. All tests were controlled internally using defined positive and negative sera. The reportable range of results for the glycoprotein ELISA was \(<0.100\) optical density (OD) units = negative, 0.100 to 0.249 OD units.
infants’ GMT with the same difference assuming an SD the same size as the preterm were seropositive for each antibody.

Appropriate. The primary outcome was antibody GMT, and the infants’ GMT, assuming an SD of 0.5 times the preterm proportion in term infants’ GMT of antibody over preterm was anticipated to be sufficient to measure a 1.5-fold eleva-
tion in GMT/H9273tion using the Mann-Whitney U test. Birth weights and ages were summarized as medians. Categorical variables were compared using the \( \chi^2 \) or Fisher’s exact tests, as appropriate. The primary outcome was antibody GMT, and the secondary outcome was the proportion of children who were seropositive for each antibody.

A patient population of 16 patients per group was anticipated to be sufficient to measure a 1.5-fold elevation in term infants’ GMT of antibody over preterm infants’ GMT, assuming an SD of 0.5 times the preterm infants’ GMT, \( \alpha = 0.05 \), and a power of 0.80, or a twofold difference assuming an SD the same size as the preterm infants’ GMT with the same \( \alpha \) and power.

**RESULTS**

Seventeen preterm infants and 19 term infants were enrolled. One preterm infant and 3 term infants had only 1 successful blood draw performed; their samples were not analyzed. Sixteen preterm infants and 16 term infants had all study procedures completed. Demographic and baseline information is summarized in Table 1.

Four patients in the preterm infant group and 1 patient in the term group had MMR and varicella vaccines given before 15 months of age; all vaccines were given between 14.4 and 16.0 months of age (Table 1). Nine preterm infants received other vaccines (9 heptavalent pneumococcal conjugate vaccine, 1 influenza vaccine) concurrent with MMR and varicella vaccines. This also occurred in 10 term infants (10 heptavalent pneumococcal conjugate vaccine, 1 influenza vaccine). Three infants in each group had minor adverse events in the 6 weeks after immunization, including low-grade fever, upper respiratory signs, and rash. A causal relationship between an adverse reaction and a specific vaccine could not be established. No serious adverse events were noted after immunization.

All patients in both groups were seronegative for measles-neutralizing antibody before vaccine, and all had measles titers \( > 120 \text{ mIU/mL} \) after vaccine (Table 2). Measles titers were similar between groups both before and after vaccine (Fig 1). Preterm infants had lower mumps and rubella titers than did term infants before vaccine; however, the groups had similar titers after vaccine (Fig 1). Similar proportions of infants had positive mumps and rubella antibody seroresponses after vaccine (Table 2). Varicella titers were similar between preterm and term infants both before and after vaccine (Fig 1). All prevaccine varicella titers in both groups fell below the limit of detection. Postvaccine varicella titers were lower than might have been expected, and several infants in both groups had titers below the defined seroresponse level (Table 2).

**DISCUSSION**

We have shown that preterm infants who were born at \(< 29 \text{ weeks’ gestation} \) mounted antibody responses to measles, mumps, rubella, and varicella after MMR and varicella vaccination at 15 months of age that were similar to those that were observed in term infants. Prevaccine mumps and rubella titers in preterm infants were lower than those in term infants. This is consistent with lower transplacental antibody transfer in the preterm infants, because the most likely source of detectable prevaccine antibodies would be maternal. Prevaccine varicella and measles titers fell near or below the assays’ limits of detection in both groups of infants, precluding any evaluation of gestational age effects.

Identification of high-risk groups for poor MMR or varicella vaccine immunogenicity potentially is clinically important. Measles, mumps, and rubella still circulate to a limited degree in the US population, with continued

---

**TABLE 1** Demographic and Baseline Information on Term and Preterm Infants Who Were Immunized at 15 Months of Age With MMR and Varicella Vaccines

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Preterm Infants (n = 16)</th>
<th>Term Infants (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestation, median (range), wk</td>
<td>26.8 (24.0–28.9)</td>
<td>40 (37–40)</td>
</tr>
<tr>
<td>Birth weight, median (range), g</td>
<td>788 (540–1250)</td>
<td>3488 (2130–4296)</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>8 (50)</td>
<td>7 (44)</td>
</tr>
<tr>
<td>White, n (%)</td>
<td>10 (63)</td>
<td>10 (63)</td>
</tr>
<tr>
<td>Mechanical ventilation in newborn period, n (%)</td>
<td>16 (100)</td>
<td>0</td>
</tr>
<tr>
<td>Systemic glucocorticoid therapy in newborn period, n (%)</td>
<td>7 (44)</td>
<td>0</td>
</tr>
<tr>
<td>Inhaled glucocorticoid therapy at time of vaccine, n (%)</td>
<td>2 (13)</td>
<td>0</td>
</tr>
<tr>
<td>Age at vaccine, median (range), mo</td>
<td>15.1 (14.9–15.9)</td>
<td>15.2 (14.4–16.0)</td>
</tr>
<tr>
<td>Vaccine to blood draw, median (range), d</td>
<td>35 (21–43)</td>
<td>32 (27–40)</td>
</tr>
</tbody>
</table>
importation from international sources. Measles and mumps cause occasional outbreaks, and these epidemics tend to cluster in populations with inadequate immunity. Measles remains a major killer of children worldwide. Varicella circulates freely in the US population and continues to cause an attenuated chickenpox illness in a significant proportion of vaccine recipients who are exposed to the virus. Individuals with lower levels of varicella immunity are more likely to experience a more severe form of varicella disease than those with higher levels of varicella immunity.

Few data existed previously for measles, mumps, rubella, or varicella vaccines that were given to infants who were born preterm. Shortly after the introduction of rubella vaccine, a study of 7 term and 5 preterm (2500 g birth weight) 12- to 25-month-old infants...
showed delayed cellular and humoral rubella vaccine responses in the preterm infants that became equivalent to those in term infants by 42 days after immunization.13

Many vaccines that first are administered under 6 months of age now have been studied fairly well in extremely preterm infants.2–13,34 Although antibody levels after vaccination in extremely preterm infants tend to be somewhat lower than those that are achieved by term infants, similar proportions of preterm and term infants mount antibody responses above the minimum levels that are considered to be protective against disease.

In contrast to our findings with MMR and varicella vaccines, limited experiences with other vaccines have suggested continued deficits beyond 6 months of age in preterm infants’ vaccine responses. Inactivated influenza vaccine, given at 6 to 18 months of age, elicits lower antibody titers and diminished T-cell proliferative responses in preterm than in term children.17 We and others have reported diminished antibody responses after tetanus, diphtheria, polio, and Haemophilus influenzae type b booster vaccines in 4- to 7-year-old children who were born extremely preterm.14–16 The immunologic mechanisms of the decreased vaccine responses are unknown but seem to affect both the humoral14–17 and the cellular17 arms of the immune system.

The relatively robust vaccine responses in the preterm infants in this study may be explained in part by the specific vaccine antigens studied. Live viral vaccines, such as MMR and varicella, often are highly immunogenic. Early studies with the varicella vaccine suggested 94% to 100% seroconversion, 93% to 100% cellular immune response, and 90% to 96% efficacy, with immunity persisting unabated over 7 to 10 years.15–38 Comparisons between oral, live, attenuated polio vaccine and injectable, enhanced-potency, inactivated polio vaccine in preterm infants also have suggested that the live vaccine may be more immunogenic.12,13,15,39

This study has inherent limitations. Relatively few infants were studied. Although the numbers were calculated to be sufficient to test the hypothesis of difference between the groups, a lack of detected difference does not necessarily suggest equivalence of the 2 study groups. However, the narrow confidence intervals that were observed in this study (Fig 1) suggest that any true difference between groups is likely to be small and clinically insignificant. Vaccines, particularly live vaccines, stimulate both the cellular and the humoral components of the immune system, and cellular responses to the vaccines studied here have been described.17,33,35,37,40 This study measured only antibody titers. Differences in antibody affinity or cellular responses between groups would not be detected. In many cases, however, antibody levels alone provide a reasonable estimate of vaccine immunogenicity.5,41–43 In addition, our results cannot rule out the possibility that some persistent immune system alteration in preterm infants may be overcome by relatively highly immunogenic, live, attenuated vaccines such as MMR and varicella vaccines.

CONCLUSION
Administration of MMR and varicella vaccine at the recommended chronologic age results in adequate antibody response to vaccine antigens, even in infants who are born before 29 weeks’ gestation. These findings support the prevailing recommendations for immunization of the preterm infant at the chronological age appropriate for a term infant.

ACKNOWLEDGMENTS
This project was funded in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (Bethesda, MD), under contract N01-A1-25460; by grant 5 M01 RR00044 from the National Center for Research Resources, National Institutes of Health, for the University of Rochester General Clinical Research Center; and by the Centers for Disease Control and Prevention.

We thank John Treanor, MD, Diane O’Brien, RN, and Doreen Francis, RN, for support of this project; Jason Roy, PhD, for statistical guidance; and the University of Rochester Neonatal Continuing Care Clinic, the Golisano Children’s Hospital at Strong Pediatric Clinic, and Elmwood Pediatrics for help with recruitment. We also thank the infants and their families for participating.

REFERENCES
8. Patel DM, Butler J, Feldman S, Graves GR, Rhodes PG. Immu-


11. Freitas da Motta MS, Mussi-Pinhata MM, Jorge SM, Tachibana Yoshiida CF, Sandoval de Souza CB. Immunogenicity of hepatitis B vaccine in preterm and full term infants vaccinated within the first week of life. Vaccine. 2002;20:1557–1562


Influenza Vaccine Coverage and Missed Opportunities Among Inner-city Children Aged 6 to 23 Months: 2000-2005

Jennifer R. Verani, MD, MPH, Matilde Irigoyen, MD, Shaofu Chen, MD, PhD, Frank Chimkin, MBA, MSW

Department of Pediatrics, Columbia University, New York, New York

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

ABSTRACT

OBJECTIVE. In 2002, the Advisory Committee on Immunization Practices recommended universal influenza vaccination of 6- to 23-month-olds. Little is known about coverage and missed opportunities for influenza vaccination at inner-city practices. The objective of this study was to assess the 2000–2001 to 2004–2005 coverage and the prevalence of missed opportunities for influenza vaccination among inner-city children.

METHODS. We conducted a retrospective review for the 2000–2001 to 2004–2005 influenza seasons at a practice network in New York City. The study population included 5 annual cohorts of 6- to 29-month olds as of March 31 of each year with ≥1 visit to the network in the previous 12 months (n = 7063). Immunization data were obtained from the network registry and the New York Citywide Immunization Registry. Coverage levels were estimated for 1 dose (partial) and 2 doses (full). Missed opportunities were assessed for visits within each influenza season.

RESULTS. Coverage rose steadily throughout the 5 years (full: 1.6% to 23.7%; partial: 1.5% to 18.1%). The relationship between year and coverage was linear. Missed opportunities occurred in 82% of visits and were more common for first (89%) than for repeat doses (38%). Missed opportunities per child per season decreased from 2.9 to 2.0 during the study period.

CONCLUSIONS. Influenza vaccine coverage among 6- to 23-month-olds at inner-city practices increased steadily from 2000–2001 through 2004–2005, and the prevalence of missed opportunities per child decreased. However, coverage remained suboptimal, with most of children not vaccinated or undervaccinated. Missed opportunities were major contributors to low coverage.
Influenza is a significant cause of morbidity for young children and infants. Hospitalization rates for influenza-attributable disease among children who are younger than 2 years are comparable to those of elderly and older children with chronic illnesses. In 2002, the Advisory Committee on Immunization Practices (ACIP) encouraged the immunization of healthy children aged 6 to 23 months, and in 2004, ACIP recommended universal vaccination of this age group. Given concerns about the viability of incorporating additional shots to an already crowded childhood immunization schedule, as well as the low coverage levels that are seen among other groups of at-risk children with long-standing recommendations for annual influenza vaccine, it is important to evaluate the uptake of this recommendation.

A small number of publications have described increases in influenza vaccine coverage levels after the changes in recommendation. Among inner-city children who were seen at a resident clinic in Nashville, Tennessee, coverage for at least 1 influenza vaccine dose increased from 1% in 2002 to 17% in 2003. In 2002–2003, on the basis of telephone surveys that were conducted nationwide, 7% of children aged 6 to 23 months had received at least 1 influenza dose; in 2004–2005, coverage had increased to 48%. Among patients who were enrolled in a large health maintenance organization in California, coverage increased from 47% in 2003–2004 to 57% in 2004–2005, based on registry data and population estimates.

Although influenza vaccine coverage is increasing, questions remain regarding the nature and the extent of the coverage and the challenges that remain. Studies have compared coverage for 1 or 2 seasons for the same population, yet descriptions of coverage that encompass the years before and after the ACIP recommendations may help understanding of patterns of uptake. Few studies have focused on inner-city populations, where immunization coverage levels are known to be low. Much of the research has relied on parental surveys, which are subject to recall bias. In addition, provider immunization delivery practices are a key determinant of vaccination levels, and additional research is needed on the contribution of provider practices to influenza vaccine coverage. For example, missed opportunities are known to contribute to low influenza vaccine coverage among children with asthma, as well as in children with chronic medical conditions, yet their impact on influenza vaccination among healthy young children is unknown. The 2 objectives of this study were (1) to assess influenza vaccine coverage levels among children aged 6 to 23 months in an inner-city population during 2000–2005, a 5-year time period that spanned the change in ACIP recommendation, and (2) to assess the impact of provider missed opportunities on influenza coverage levels in this age group.

Methods

Study Setting

The study was conducted at a network of 5 community-based pediatric practices that are affiliated with an academic health center at an inner-city community in New York City. The network is the major pediatric health care provider in the community (>63,000 pediatric visits in 2005, ranging from 8000 to 20,000 per practice). The patient population predominantly is Latino and black. Most children are covered by Medicaid, and of these, half are enrolled on managed care plans. The Vaccine for Children Program provides the majority of vaccines (>90%) that are given at the practices.

Study Design and Population

We conducted a retrospective cohort study of influenza vaccine coverage among children who were aged 6 to 23 months for 5 consecutive influenza seasons. The eligibility criteria included (1) age 6 to 29 months as of March 31 (2001–2005) of each year and (2) at least 1 visit to the network in the previous 12 months. We included children who were aged up to 29 months because they would have been ≥23 months at the start of the season and therefore eligible for vaccination. The children then were grouped into 5 annual cohorts of children, 2000–2001 through 2004–2005, with a cohort sample size ranging from 1396 to 1428. Each child could have been included in up to 2 cohorts.

Data Sources

Children were identified through the network’s billing/registration system, which was the source for demographic and visit data. Children were classified as Medicaid participants when they were currently or at any time in the past insured by Medicaid.

Vaccine data were collected from both the network immunization registry, EzVAC, and the New York City-wide Immunization Registry (CIR). EzVAC is a point-of-service registry that is linked to the hospital billing system and includes all children who receive care at the hospital or affiliated network. CIR is New York City’s population-based registry. EzVAC has a capture proportion of 95% of vaccines administered at network sites (M. Irigoyen, MD, verbal communication, 2006). New York City providers are mandated to report to CIR: >90% of providers report to CIR, and, of these, 80% do so regularly (A. Metroka, MSW, written communication, 2003).

Outcome Measures

For each cohort, we determined influenza vaccine coverage levels. We defined “full coverage” as 2 doses of influenza vaccine, either 2 doses in 1 season or 1 dose during the season of interest and at least 1 dose during the previous season. We defined “partial coverage” as...
Characteristic children were insured by Medicaid. There were no significant differences across the 5 cohorts with regard to age, gender, or insurance type. The annual cohorts ranged in size from 1396 to 1428 with a mean of 1413 and a total N = 7063 for the 5-year period (Table 2). A total of 88.4% of vaccines were recorded in both registries: 5.2% of vaccines were recorded in EzVac only, and 6.5% were recorded in CIR only. Each season, >90% of vaccines were administered between October and January (Fig 1).

**RESULTS**

**Study Population**

As shown in Table 1, half of the children were male, and there was an even distribution across age groups. Most children were insured by Medicaid. There were no significant differences across the 5 cohorts with regard to age, gender, or insurance type. The annual cohorts ranged in size from 1396 to 1428 with a mean of 1413 and a total N = 7063 for the 5-year period (Table 2). A total of 88.4% of vaccines were recorded in both registries: 5.2% of vaccines were recorded in EzVac only, and 6.5% were recorded in CIR only. Each season, >90% of vaccines were administered between October and January (Fig 1).

**Coverage Levels**

Coverage levels increased throughout the 5-year period (Table 2). From 2000 though 2005, coverage for any dose rose from 3.2% (95% CI: 2.2%–4.1%) to 41.8% (95% CI: 39.2%–44.3%). The proportion of children who were fully immunized increased from 1.6% (95% CI: 1.0%–2.3%) to 23.7% (95% CI: 21.5%–25.9%) during the 5-year period, and the proportion who were partially vaccinated (required 2 doses but received only 1) increased from 1.5% (95% CI: 0.9%–2.2%) to 18.1% (95% CI: 16%–20.1%). The relationship between year and coverage was linear for both fully and partially immunized (R² = 0.9139 and R² = 0.9322, respectively; Fig 2).

Coverage initially was higher among the children aged 25 to 29 months at the end of the season. However, in the years after the change in recommendation, coverage increased most dramatically among 13- to 24-month-olds. In 2004–2005, coverage was highest among those aged 13 to 18 months (31.5% [95% CI: 26.7%–36.4%] were fully vaccinated) and lowest among those aged 6 to 12 months (16.9% [95% CI: 13.0%–20.7%]; Fig 3).

**Missed Opportunities**

Missed opportunities occurred in 82.2% (95% CI: 81.5%–83.0%) of all vaccine-eligible visits. Of the visits during which a first dose of vaccine could have been administered, 89.2% (95% CI: 88.6%–89.8%) resulted in missed opportunities. Among visits that were eligible for a second dose, the proportion of missed opportunities was lower (38.4% [95% CI: 35.9%–40.8%]). The months with the greatest proportion of missed opportunities varied from year to year, with no clear trend over the 5-year study period.

When looking at missed opportunities that were experienced by each child, we found that the proportion of children in each cohort with at least 1 missed opportunity ranged from a high of 61% (95% CI: 58.8%–63.9%) in 2001–2002 to a low of 40% (95% CI: 37.5%–42.7%) in 2004–2005 (Fig 4). The average number of missed opportunities per child decreased from 2.9 in 2000–2001 to 2.0 in 2004–2005 (Fig 5). The proportion of all visits with missed opportunities decreased steadily
during the 5-year period—from 97% in 2000–2001 to 55% in 2004–2005—and was inversely proportional to coverage level (Fig 6).

### DISCUSSION

In response to a growing body of evidence that young children are at high risk for significant influenza-related morbidity, in 2002, ACIP expanded its recommendations to include vaccination of children aged 6 to 23 months. This study found that in the 5-year period that spanned the change in recommendation, coverage levels for 6- to 23-month-olds at an inner city practice network increased steadily. Nonetheless, even in 2004–2005, more than three quarters of children were either nonvaccinated or undervaccinated. Each year, nearly half of children who received 1 dose of vaccine failed to receive a required second dose. These low coverage lev-

### TABLE 2: Influenza Vaccine Coverage Levels: Children Aged 6 to 23 Months, 2000–2005

<table>
<thead>
<tr>
<th>Influenza Season</th>
<th>n (N = 7063)</th>
<th>Any Coverage, % (95% CI)</th>
<th>Full Coverage, % (95% CI)</th>
<th>Partial Coverage, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000–2001</td>
<td>1428</td>
<td>3.2 (2.2–4.1)</td>
<td>1.6 (1.0–2.3)</td>
<td>1.5 (0.9–2.2)</td>
</tr>
<tr>
<td>2001–2002</td>
<td>1420</td>
<td>4.5 (3.4–5.6)</td>
<td>2.3 (1.5–3.1)</td>
<td>2.2 (1.4–2.9)</td>
</tr>
<tr>
<td>2002–2003</td>
<td>1416</td>
<td>14.9 (13.0–16.8)</td>
<td>8.1 (6.7–9.5)</td>
<td>6.8 (5.5–8.1)</td>
</tr>
<tr>
<td>2003–2004</td>
<td>1403</td>
<td>23.7 (21.5–26.0)</td>
<td>13.0 (11.3–14.8)</td>
<td>10.7 (9.1–12.3)</td>
</tr>
<tr>
<td>2004–2005</td>
<td>1396</td>
<td>41.8 (39.2–44.3)</td>
<td>23.7 (21.5–25.9)</td>
<td>18.1 (16.0–20.1)</td>
</tr>
</tbody>
</table>
els are striking when compared with the high levels of coverage that were seen for other childhood vaccines in the United States.21,22 The range of full and partial influenza vaccine coverage levels in this population falls within those reported by previous studies.17–19,25–28 The increase in coverage after the change in recommendation also is consistent with previously published data. The perspective of a 5-year window provides additional insight into the adaptation of the new policy. One might expect a marked increase in coverage level in the season immediately after the change in recommendation. However, our data showed a more gradual, linear increase in coverage from 2000–2005, with suboptimal coverage even at the end of the 5-year period.

Our study found extremely high numbers of missed opportunities for influenza vaccination: >80% of all possible occasions to vaccinate were missed. The proportion of visits with missed opportunities per season decreased markedly during the 5-year study period, and the inverse relation with coverage levels highlights the crucial role that missed opportunities play in determining influenza immunization coverage. Missed opportunities have been shown to contribute to low levels of influenza vaccination among children with asthma23 and other high-risk conditions.24 Notably, a much larger proportion of first dose–eligible visits resulted in missed opportunities as compared with second dose–eligible visits, suggesting fewer barriers to receiving a second dose once a first dose of vaccine has been administered. Our data do not provide insight into the reasons behind missed opportunities (eg, oversight, physician noncompliance, parental opposition). More research in this area is required, and future efforts to improve vaccination coverage should focus on reducing missed opportunities, particularly for the first dose.

Low coverage and missed opportunities may reflect a lack of physician compliance with the new recommendation. According to the awareness-to-adherence model, physicians progress through 4 phases in the process of changing clinical behaviors in response to clinical guidelines: awareness, agreement, adoption, and adherence.29 Awareness of influenza vaccine guidelines may have been affected by the changing nature of the recommendation. In 2002 and 2003, ACIP urged physicians to “encourage when feasible” the vaccination of healthy children aged 6 to 23 months,9 but in 2004, ACIP “recommended” universal vaccination in this age group.10 A survey that was conducted in 2001 reported that only 50% of pediatricians and 40% of family practitioners believed that universal influenza vaccination of children aged 6 to 23 months would be feasible.30 The perception that one is not capable of implementing a new guideline is an important barrier to physician compliance.31

Parental beliefs and attitudes also can serve as a barrier to immunization. Parents frequently have concerns about the safety of the influenza vaccine,32,33 particularly inner-city parents.34 Parents often do not perceive their child to be at risk.33,35 Parental attitudes may have changed after the 2003–2004 season, which received substantial media attention regarding severe influenza cases among children.36 Surveys of parents consistently find that the recollection of a physician recommendation for influenza vaccine is associated with higher reported vaccination levels.35–37

Studies have demonstrated the effectiveness of interventions that aim to increase influenza vaccine coverage among 6- to 23-month-olds. A multifaceted strategy that comprises education, reminders, standing orders, and express immunization service had a strong positive impact at 10 inner-city clinics in Pittsburgh, Pennsylvania, during the 2002–2003 season.27 Likewise, a registry reminder intervention proved effective at increasing coverage during the 2003–2004 season in Denver, Colorado.38 There is significant evidence to support the use of
provider reminder recall and provider assessment and feedback to reduce missed opportunities for vaccination among children. 

Given the low levels of vaccination among children aged 6 to 23 months found in this and other studies, such interventions will be needed to achieve adequate immunization coverage.

This study had several limitations. The coverage levels may have been underestimated as a result of underreporting to 1 or both registries. This problem may be compounded by use of multiple providers and record scatter. 

The study population was limited to patients of 1 practice network that serves inner-city children, and the findings may not be generalizable to other types of practices or patient populations. We assumed a steady supply of vaccine starting from the first date of administration of an influenza vaccine at each site; however, shortages may have affected delivery of vaccine and coverage. Although shortages theoretically should not have affected coverage in this prioritized at-risk group, there were reports in 2004–2005 of difficulties in obtaining vaccine for children aged 6 to 23 months, and it is unclear how shortages may have affected coverage during the 5-year period. Last, we may have overestimated missed opportunities; some of these may be attributable to parent refusals.

CONCLUSIONS

Influenza vaccine coverage among a population of inner-city children aged 6 to 23 months increased steadily from 2000 through 2005. However, even in the final year of analysis, coverage was suboptimal, with more than three quarters of children in this age group remaining nonvaccinated or undervaccinated. Throughout the study period, >80% of all opportunities to vaccinate were missed. The proportion of visits with missed opportunities by season decreased during the 5-year period and was inversely proportional to coverage levels. In the future, continuous monitoring of coverage among this age group is required. It is important to investigate factors that contribute to missed opportunities, particularly for the first dose. The role of vaccine supply and shortages on coverage in this age group is another area for future research. Interventions that are known to be effective at increasing influenza vaccine coverage among other high-risk groups should be expanded to include this age group.

REFERENCES

24. Daley MF, Beaty BL, Barrow J, et al. Missed opportunities for
Vaccine Effectiveness Against Medically Attended, Laboratory-Confirmed Influenza Among Children Aged 6 to 59 Months, 2003–2004

Carrie M. Shuler, DVM, MPH\textsuperscript{a,b}, Martha Iwamoto, MD, MPH\textsuperscript{c}, Carolyn Buxton Bridges, MD\textsuperscript{d}, Mona Marin, MD\textsuperscript{a}, Ruth Neeman, RN, BSN\textsuperscript{b}, Paul Gargiullo, PhD\textsuperscript{c}, Terrace A. Yoder, MD\textsuperscript{d}, Harry L. Keyserling, MD\textsuperscript{c}, Pauline D. Terebuh, MD, MPH\textsuperscript{b}

\textsuperscript{a}Epidemic Intelligence Service, \textsuperscript{b}Centers for Disease Control and Prevention, Atlanta, Georgia; \textsuperscript{c}Notifiable Diseases Section, Division of Public Health, Georgia Department of Human Resources, Atlanta, Georgia; \textsuperscript{d}Children’s Medical Group, Atlanta, Georgia; \textsuperscript{e}School of Medicine, Emory University, Atlanta, Georgia

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

ABSTRACT

OBJECTIVES. Influenza is a leading cause of illness among children. Studies rarely have measured influenza vaccine effectiveness among young children, particularly when antigenic match between vaccine and circulating viruses is suboptimal. We assessed vaccine effectiveness against medically attended, laboratory-confirmed influenza for children who were aged 6 to 59 months during the 2003–2004 influenza season.

METHODS. In a case-control study that was conducted in a single pediatric practice, case patients who were aged 6 to 59 months and had laboratory-confirmed influenza were age matched 1:2 to eligible control subjects. Vaccination status was ascertained as of the date of the case patient’s symptom onset. Conditional logistic regression was used to calculate vaccine effectiveness, adjusting for underlying medical conditions and health care usage.

RESULTS. We identified 290 influenza case patients who were seen for medical care from November 1, 2003, to January 31, 2004. Vaccine effectiveness among fully vaccinated children, compared with unvaccinated children, was 49%. Partially vaccinated children who were aged 6 to 23 months had no significant reduction in influenza (vaccine effectiveness: \textsuperscript{70%), but partially vaccinated children who were aged 24 to 59 months had a significant (65%) reduction in influenza, compared with unvaccinated children.

CONCLUSIONS. Full vaccination provided measurable protection against laboratory-confirmed influenza among children who were aged 6 to 59 months during a season with suboptimal vaccine match. No vaccine effectiveness was identified with partial vaccination among children who were aged 6 to 23 months, affirming that children need to be fully vaccinated to obtain protective effects. These results strengthen the evidence of the vaccine’s ability to reduce substantially the burden of disease in this age group.
Influenza is one of a limited number of viral respiratory diseases that are preventable by vaccination. It is common among young children, resulting in an estimated annual average of 50 to 95 outpatient visits and 6 to 27 emergency department visits per 1000 children, 9 hospitalizations per 10 000 children, and an annual average of ~92 deaths among US children aged <5 years.1–4 Compared with older children, children who are aged <24 months have higher rates of acute otitis media, pneumonia and other lower respiratory tract disease, and hospitalization related to the influenza infection.1,5–8

For maximization of the immune response and protection against influenza, previously unvaccinated children who are aged <9 years and for whom vaccine is recommended should receive 2 doses of influenza vaccine at least 1 month apart.9 The short period between vaccine availability and the start of the influenza season, however, can make receipt of 2 doses before influenza exposure difficult and requires extra health care provider visits.10–12

Limited data have been published on the effectiveness of influenza vaccine among young children overall and particularly on the effectiveness of 1 vs 2 doses.13–17 Such data are especially important during years of influenza vaccine shortage or when the influenza season occurs early, because children might not have received their second dose of vaccine before being exposed.

For the 2003–2004 influenza season, the Advisory Committee on Immunization Practices (ACIP) encouraged vaccinating all children who are aged 6 to 23 months and household contacts of children aged <2 years and continued to strongly recommend the vaccination of children who are aged ≥6 months with medical conditions that place them at increased risk for influenza-related complications.18 In 2004, ACIP recommended annual vaccination for all children who are aged 6 to 23 months.19

Nationally, during the 2003–2004 influenza season, influenza viruses began circulating unusually early and influenza A (H3N2) viruses predominated.20 In Georgia, the first reported laboratory-confirmed patient had illness onset in October, and overall influenza activity peaked in late December.21 Only 25% of the circulating influenza viruses nationally and in Georgia were similar antigenically to the vaccine strain, and severe complications, including pediatric deaths, were reported.20,22 Because clinical illness that is caused by influenza virus is not distinguishable from other causes of respiratory illness, laboratory confirmation of influenza is essential to increase the accuracy of vaccine effectiveness (VE) estimates. Use of nonspecific outcomes for influenza diagnoses may lower VE estimates substantially.23 We evaluated the effectiveness of the trivalent inactivated influenza vaccine against medically attended, laboratory-confirmed influenza among children who were aged 6 to 59 months in Georgia during a season with a suboptimal antigenic match between the vaccine and circulating strains.

Methods

Study Population and Location

Children who were examined for 1 or more outpatient visits (either well-child or sick visits) at a private metropolitan Atlanta pediatric practice during February 1, 2003, to January 31, 2004, and who were born during October 1, 1998, to April 1, 2003 (ie, aged 6–59 months during the 2003–2004 influenza season), were eligible for the study. We identified 7139 children who met the inclusion criteria (study children). The study practice has 16 health care providers and serves ~400 patients a day. The physicians offered influenza vaccine to patients consistent with the 2003–2004 ACIP guidelines (children aged ≥6 months at high-risk, all children aged 6–23 months, and household contacts of children aged <2 years) and routinely performed in-house rapid influenza antigen testing on children with influenza-like illness (ILI). This assessment of influenza VE was initiated as a public health response; therefore, the Centers for Disease Control and Prevention and the Georgia Division of Public Health determined these activities as nonresearch and did not require review by an institutional review board. However, personal identifiers were removed and confidentiality was preserved in data collection and analysis.

Laboratory-Confirmed Influenza Outcome

The practice provided the QuickVue Influenza test (Quidel Corp, San Diego, CA) to screen nasopharyngeal swab samples for influenza antigen. This rapid antigen test has a sensitivity of ~73% and specificity of ~96%.24 Because of a test kit shortage in the community, additional kits were provided to the practice on January 5, 2004, to allow for influenza testing at the practice without interruption. To determine the feasibility of conducting the retrospective observational study at the pediatric practice, a survey of the providers was conducted to evaluate potential diagnostic testing biases. The providers were asked how their patients’ vaccination status, age, underlying medical conditions, clinical symptoms, and insurance reimbursement affected their decision to test for influenza (more likely to test, less likely to test, or would not affect decision to test).

Case Patient Identification

Case patients were identified as those with laboratory-confirmed influenza during November 1, 2003, to January 31, 2004. We conducted case finding through review of the study practice’s laboratory logbooks. Case patient medical charts were reviewed by using a chart-abstraction tool. Chart abstractors were not blinded to
the patient’s case status. Information was collected on illness symptoms, treatment with influenza antiviral medication, influenza complications, influenza vaccination, gender, chronic medical conditions, and child care or school participation.

Control Subject Selection
Two age-matched study children were randomly selected as control subjects for each case patient. The eligible control subjects for each case patient included study children who were born during the same month and year and had not become a case before the influenza illness onset date of the case patient. We excluded from the study non-Georgia residents and children with no vaccination history indicated in the medical chart. Case patients were eligible to serve as control subjects until the date of chart-documented, laboratory-confirmed influenza. Although limited in-and-out migration from the study population probably occurred, any child with a visit to the practice during the previous year was considered at risk during the influenza season and became unavailable for control section only on the date of becoming a case patient and thereafter. We did not exclude children who were examined at the practice with ILI, who were not tested for influenza. Selecting a child as a control subject more than once for case patients with different symptom onset dates did not necessarily lead to repeat exposure information, because vaccination status varied with time. The vaccination status of each set of case patients and control subjects was compared on the date of illness onset of the case patient.

Vaccination Status
By using the date of symptom onset as the anchor date, vaccination status and other covariates for case patients and their age-matched control subjects were categorized. We considered children who had received 2 doses of influenza vaccine at least 1 month apart and ≥14 days before the anchor date as fully vaccinated. Children who were vaccinated during a previous season needed only 1 dose of vaccine during the 2003–2004 season ≥14 days before the anchor date to be fully vaccinated. We categorized children with the following 2 situations as partially vaccinated: (1) children who were not vaccinated in a previous season and had received 2 doses of influenza vaccine since September 2003 with an anchor date <14 days after the second dose and (2) children who were not vaccinated in a previous season and received only 1 dose of vaccine since September 2003 ≥14 days before the anchor date. We considered children who had received no doses of influenza vaccine during the 2003–2004 season on or before the anchor date and children who had received 1 dose since September 2003 <14 days before the anchor date as unvaccinated, even if they had received vaccine during a previous season.

Presence of Underlying Medical Conditions
ACIP recommendations for multiple years have targeted children with certain medical conditions to be at increased risk for complications related to influenza virus infection. We sought documentation of high-risk conditions in the medical chart and included conditions that were designated by the ACIP during the 2003–2004 season: reactive airway disease (RAD) or asthma (listed explicitly on at least 2 occasions), chronic lung disease, chronic metabolic disease, cardiovascular disease, renal disease, hemoglobinopathy, cancer, other immunosuppressive conditions, long-term aspirin therapy, or residence in a long-term care facility.

Child Care or School Status
We ascertained child care or school participation to adjust for the possible increase in influenza virus exposure; priming from previous influenza infections has been reported to be associated with higher immune response after vaccination. The practice physicians collected on a standardized form information on child care or school attendance during health care visits. When no notation had been made in the medical chart during our study period or the section on the medical chart had been left blank, we recorded the child’s out-of-home care status as unknown.

Health Care Usage
We used the number of office visits billed for each child during the year before the beginning of the 2003–2004 flu season (November 1, 2002, to November 1, 2003) as an indicator of health care usage. When the number of visits was more than the median value for the child’s corresponding age category, the child was classified as a high health care user.

Statistical Analysis
We modeled the association between influenza vaccination and medically attended, laboratory-confirmed influenza with conditional logistic regression. A matched analysis was used to estimate an odds ratio (OR) for disease in which each matched set (1 case patient and 2 control subjects) was treated as a unique stratum. We included high-risk status and health care usage in the final models. We analyzed vaccination status as a categorical exposure measure with unvaccinated as referent. Fully vaccinated children were compared with unvaccinated children. Partially vaccinated children were compared with unvaccinated children who might have become partially vaccinated during the 2003–2004 influenza season (ie, children who were categorized as unvaccinated and were not vaccinated during a previous season). We also calculated separate VE estimates for children who were aged 6 to 23 months and 24 to 59 months. The adjusted OR (aOR) was used to estimate VE by using the formula VE = (1 − aOR) × 100.
RESULTS

Provider Testing Practices

Of the 16 providers, 8 completed the survey. One provider reported that patients’ vaccination status affected the decision to test for influenza (more likely to test unvaccinated patients and less likely to test vaccinated patients). The other respondents reported that patients’ vaccination status did not affect their decision to test. Many of the 8 practitioner-respondents identified other clinical factors that would make them more likely to test a patient for influenza: patient high-risk condition (n = 7), household contact at increased risk for an influenza-related complication (n = 8), fever >104°F (n = 6), illness duration <2 days (n = 4), and decision to prescribe antiviral medications (n = 8). A fever <101°F was identified by 5 respondents as a patient factor that would make them less likely to test.

Description of Case Patients

A total of 293 children who were aged 6 to 59 months and had laboratory-confirmed influenza during November 1 to January 31, 2004, were identified from the practice’s records. Three children did not have complete vaccination records and were excluded from the analysis (N = 290; Fig 1). The majority of case patients who were identified early in the season were unvaccinated. The epidemiologic curve of case patients who were identified in the practice was similar to the curve that was identified through Georgia’s influenza surveillance.

The case patients’ clinical course and treatment are described in Table 1 by age category (6–23 months and 24–59 months). Fever, cough, and rhinorrhea were listed as symptoms for 97%, 89%, and 83% of all case patients, respectively. The physicians prescribed oseltamivir for 133 (46%) of the influenza case patients, and 1 patient was prescribed amantadine. Influenza-related complications were identified within 3 weeks of their influenza illness onset among 32% of children who were aged 6 to 23 months and 22% of children who were aged 24 to 59 months (Table 1). Only 2 children were hospitalized as a result of their influenza infection: a child who was aged 16 months and had a history of RAD/asthma and a child who was aged 9 months and was not at high risk. Both children were fully vaccinated for influenza at the time of illness onset.

A comparison of influenza case patient symptoms before and after the receipt of supplemental test kits to the practice found that children who were aged 6 to 23 months and tested positive for influenza were less likely to present with a cough, and children who were aged 24–59 months who tested positive for influenza were less likely to present with rhinorrhea.
to 59 months were less likely to present with vomiting after the test kits were provided compared with before. Otherwise, there was no difference in the presenting symptoms of influenza-positive patients before and after the test kits were provided to the practice.

Characteristics of Case Patients and Age-Matched Control Subjects

Table 2 describes the vaccination status, gender, medical conditions, child care participation, and health care usage of the case patients and their age-matched control subjects. A higher proportion of children who were aged 6 to 23 months (42.7% of the case patients and 56.8% of control subjects) compared with older children were considered fully vaccinated at the time of their respective anchor dates. Among children who were aged 24 to 59 months, 22.3% of case patients and 30.5% of control subjects were fully vaccinated. Gender was evenly distributed between the age categories and between the case patients and control subjects.

<table>
<thead>
<tr>
<th>Characteristics, n (%)</th>
<th>Aged 6–23 mo</th>
<th>Aged 24–59 mo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case Patients (n = 111)</td>
<td>Control Subjects (n = 222)</td>
</tr>
<tr>
<td>Vaccination status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fully vaccinated</td>
<td>47 (42.7)</td>
<td>126 (56.8)</td>
</tr>
<tr>
<td>2 doses</td>
<td>25 (22.5)</td>
<td>66 (29.7)</td>
</tr>
<tr>
<td>1 dose and vaccinated in previous season</td>
<td>22 (19.8)</td>
<td>60 (27.0)</td>
</tr>
<tr>
<td>Partially vaccinated</td>
<td>26 (23.5)</td>
<td>57 (26.5)</td>
</tr>
<tr>
<td>2 doses</td>
<td>4 (3.6)</td>
<td>5 (2.3)</td>
</tr>
<tr>
<td>1 dose</td>
<td>21 (18.9)</td>
<td>22 (9.9)</td>
</tr>
<tr>
<td>Unvaccinated</td>
<td>39 (35.1)</td>
<td>69 (31.0)</td>
</tr>
<tr>
<td>Vaccinated in previous season</td>
<td>9 (8.1)</td>
<td>14 (6.3)</td>
</tr>
<tr>
<td>Not vaccinated in previous season</td>
<td>30 (27.1)</td>
<td>55 (24.8)</td>
</tr>
<tr>
<td>Vaccinated in previous season</td>
<td>31 (27.9)</td>
<td>78 (35.1)</td>
</tr>
<tr>
<td>Male gender</td>
<td>55 (49.5)</td>
<td>119 (53.9)</td>
</tr>
<tr>
<td>High-risk conditions&lt;sup&gt;a,d&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancer</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>1 (0.9)</td>
<td>5 (2.3)</td>
</tr>
<tr>
<td>RAD/asthma</td>
<td>22 (19.8)</td>
<td>33 (14.9)</td>
</tr>
<tr>
<td>Renal disease</td>
<td>1 (0.9)</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Chronic lung disease</td>
<td>0 (0)</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Immunosuppression</td>
<td>0 (0)</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Metabolic disease</td>
<td>0 (0)</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Other</td>
<td>0 (0)</td>
<td>2 (0.9)</td>
</tr>
<tr>
<td>Any</td>
<td>24 (21.6)</td>
<td>44 (19.8)</td>
</tr>
<tr>
<td>Child care/school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>58 (52.5)</td>
<td>87 (39.2)</td>
</tr>
<tr>
<td>No</td>
<td>53 (47.8)</td>
<td>134 (60.4)</td>
</tr>
<tr>
<td>Unknown</td>
<td>0 (0)</td>
<td>1 (0.4)</td>
</tr>
<tr>
<td>No. of office visits</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median visits</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>High health care usage&lt;sup&gt;a&lt;/sup&gt;</td>
<td>71 (64.0)</td>
<td>110 (49.6)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Not vaccinated in a previous season.<br><sup>b</sup> Second dose <14 days before symptom onset.<br><sup>c</sup> First dose ≥14 days before symptom onset.<br><sup>d</sup> Case patients and control subjects might have had >1 medical condition.<br><sup>e</sup> Above the median number of office visits for the year before the 2003–2004 influenza season for the age category: 6 to 59 months = 6.0; 6 to 23 months = 10.0; 24 to 59 months = 4.0.
VE Estimates
Vaccination status, health care usage, and child care participation were associated significantly with influenza in the univariate model. However, in the multivariate model, only vaccination status and health care usage were associated independently with laboratory-confirmed influenza. VE estimates were adjusted for health care usage, an important confounder, and high-risk status for comparability with previously reported VE estimates (Table 3). In the multivariate model, child care status did not affect the relationship between vaccination and having influenza and was not included in the final model. For children who were aged 6 to 59 months, fully vaccinated children had a significant reduction in influenza when compared with unvaccinated children, with a VE estimate of 49% (95% confidence interval [CI]: 30%–60%). Partially vaccinated children who were aged 6 to 59 months did not have a significant reduction in influenza (aOR: 0.76; 95% CI: 0.5–1.2).

Fully vaccinated children who were aged 6 to 23 months had a significant (52%) reduction in influenza, compared with unvaccinated children (95% CI: 20%–70%). Children who were aged 6 to 23 months and were partially vaccinated did not have a significant reduction in influenza, compared with unvaccinated children who were not vaccinated in a previous season (aOR: 1.70; 95% CI: 0.9–3.8). A significant (45%) reduction in influenza for children who were aged 24 to 59 months was identified when fully vaccinated children were compared with unvaccinated children (95% CI: 10%–70%). Partially vaccinated children who were aged 24 to 59 months had a significant (65%) reduction in influenza when compared with unvaccinated children who had not been vaccinated during a previous season (95% CI: 30%–80%).

DISCUSSION
Fully vaccinated children who were aged 6 to 59 months and partially vaccinated children who were aged 24 to 59 months had significant reductions in medically attended, laboratory-confirmed influenza, compared with respective unvaccinated children, even during a year with a suboptimal match between the vaccine and circulating strains. However, we identified no VE with partial vaccination among children who were aged 6 to 23 months. These results support recommendations for annual influenza vaccination of children and highlight the importance of achieving full vaccination before the start of the influenza season.

Comparison With Other Studies
Our results are consistent with other studies that have evaluated either the efficacy or the effectiveness of the influenza vaccine among children through randomized, controlled trials (efficacy) and observational studies (effectiveness). VE estimates can vary widely, depending on the study population (age, immune health, previous exposure to influenza, and access to vaccine), specificity of the outcome (laboratory-confirmed influenza through serology, culture, or rapid antigen tests or such symptoms as ILI), and the influenza season (how the vaccine strains relate to the circulating strains and relative contribution of influenza to the burden of ILI).

In a study by Ritzwoller et al during the 2003–2004 influenza season, VE against ILI (not laboratory-confirmed) for fully vaccinated children who were aged 6 months to 8 years was 23% and 51% for pneumonia and influenza (P&I). For fully vaccinated children who were aged 6 to 23 months, VE was 25% against ILI and 49% against P&I. VE against P&I yielded a similar VE to the laboratory-confirmed influenza outcome in our study. Unlike the study by Ritzwoller et al, our study identified a protective effect for partial vaccination among children who were aged 24 to 59 months. This difference might be caused by our population’s having had more previous exposures to influenza, or the more specific outcome of laboratory-confirmed influenza allowed for the detection of this benefit. Similar to Ritzwoller et al, our study provides reassurance that vaccination of young children provides benefit, even in a year with a suboptimal match.

A limited number of studies have been published on the effectiveness or efficacy of the influenza vaccine among children aged who are <5 years. In a Cochrane review of studies that were conducted among healthy children, published in early 2005, only 1 study

| TABLE 3 Multivariate Conditional Logistic Regression Analysis of Laboratory-Confirmed Influenza by Vaccination Status, Adjusting for High-Risk Status and Health Care Usage According to Age Category (N = 870) |
|----------------------------------|-----------------|-----------------|-----------------|
| Parameter                        | Aged 6–59 mo aOR (95% CI) VE, % | Aged 6–23 mo aOR (95% CI) VE, % | Aged 24–59 mo aOR (95% CI) VE, % |
| Fully vaccinated versus unvaccinated | 0.51 (0.4–0.7) a 49 | 0.48 (0.3–0.8) a 52 | 0.55 (0.3–0.9) a 45 |
| Partially vaccinated versus unvaccinated | 0.76 (0.5–1.2) | 1.7 (0.9–3.8) | 0.35 (0.2–0.7) a 65 |
| High health care usage | 1.6 (1.1–2.1) a | 2.5 (1.4–4.2) a | 1.2 (0.8–1.8) |
| High-risk conditions | 0.9 (0.7–1.3) | 1.0 (0.6–1.8) | 0.9 (0.6–1.3) |

a P < 0.1; b P < 0.05; otherwise not significant.

c Unvaccinated children not vaccinated during a previous season.

d Above the median number of office visits for the age category: 6 to 59 months = 6.0, 6 to 23 months = 10.0, 24 to 59 months = 4.0.
of children who were aged ≤2 years by Hoberman et al\textsuperscript{14} was included. In that 2-year randomized, placebo-controlled study, vaccine efficacy against culture-confirmed influenza was 66\% during a year when the attack rate was 16\% among unvaccinated children and ~7\% during a second year, when only 3\% of children had influenza. A limited number of children during the second-year study cohort were influenza culture positive, substantially reducing the ability to detect a difference between the vaccine groups. The decreased sensitivity and specificity of the rapid antigen test that was used in our study compared with culture and the less optimal antigenic match between the vaccine and circulating strains might explain the lower VE that was identified during our study.

Our study was unique in that the study population was highly vaccinated: 77\% of study participants who were aged 6 to 23 months had ≥1 dose of influenza vaccine by the end of the season. In Georgia overall, only 19.5\% (95\% CI: 12.9–28.4) of children who were aged 6 to 23 months during the 2003–2004 influenza season had had ≥1 dose of influenza vaccine.\textsuperscript{33} Providers at the pediatric practice routinely tested patients with suspicion of influenza. For young children who might experience numerous medically attended respiratory illnesses during the first years of life, our main outcome measure of medically attended, laboratory-confirmed influenza rather than ILI is a substantial advantage in providing an influenza-specific VE estimate. In addition, chart and billing record review of all case patients and control subjects permitted precise ascertainment of influenza vaccination dates, underlying medical conditions, health care usage, and child care or school attendance. These factors can influence a patient’s vaccination status (opportunity or health indication for vaccination) and also are related to diagnosis of laboratory-confirmed influenza. Adjusting for these variables allowed us to control for their potentially confounding effect in our assessment of influenza VE.\textsuperscript{31,34}

**Partially Vaccinated Children**

We found a significant reduction in laboratory-confirmed influenza among partially vaccinated children who were aged 24 to 59 months. The VE estimate for partially vaccinated children who were aged 24 to 59 months was higher than that for fully vaccinated children, but the sample sizes were small and the CIs for the 2 estimates overlapped. Older children are less likely to be immunologically naïve and therefore can produce a protective response with a single dose of influenza vaccine.\textsuperscript{28} However, given the yearly variability in the circulation of influenza viruses from community to community, many children who are aged 24 months to 8 years are likely not to have been exposed to circulating influenza viruses.\textsuperscript{28,35} Therefore, previously unvaccinated children who are aged <9 years should continue to be provided 2 doses to offer maximum protection from circulating types and subtypes of influenza.

**Limitations**

Observational VE studies similar to our study have intrinsic limitations.\textsuperscript{23} Bias in selection of patients for vaccination or in the diagnosis of influenza cannot be excluded because of variability in patient health care-seeking behaviors. Providers were aware of the vaccination status of the children when deciding which patients to test for influenza. Preferentially vaccinating patients with an increased risk for exposure to influenza (eg, children who were enrolled in child care) might decrease the VE estimate, whereas preferentially testing unvaccinated patients might increase the VE estimate. We attempted to minimize these potential confounding effects by evaluating the contribution of high-risk status and child care participation, by controlling for health care usage, and by choosing a pediatric practice study site that strongly supported the ACIP vaccination recommendations and routinely tested patients who had ILI for influenza. The provider survey indicated that most of the practitioners were not influenced by vaccination status when deciding to test for influenza; however, a standardized protocol for influenza testing did not exist, and testing was at the discretion of individual providers.

Although using influenza rapid testing for laboratory confirmation of medically attended influenza enhanced the precision of the influenza outcome measure, misclassification of case patients and control subjects still was possible. Patients with mild or atypical influenza symptoms might not have been tested for influenza, although we have no reason to believe that symptoms differed among vaccinated compared with unvaccinated children with influenza. Virus shedding is greatest during the first days of influenza infection; therefore, patients who were tested during the early phase of the illness were more likely to test positive. Because the influenza rapid test is ~73\% sensitive, certain medically attended cases of influenza might not have been detected. With a rapid test specificity of 96\%, certain patients might have falsely tested positive. Overall, the potential to misclassify case patients as control subjects was greater and might have decreased the VE estimate. Because the study population became more vaccinated as the season progressed and additional rapid test kits were made available to the practice, the impact of detecting milder cases later in the season on VE estimates is difficult to predict. It is unlikely to have been profound, because individual providers had established approaches to evaluating ILI and testing for influenza. Overall, most factors that could affect the VE estimate would result in an underestimation of the vaccine’s true effectiveness.
CONCLUSIONS
Influenza vaccination provided protection against medically attended, laboratory-confirmed influenza among children who were aged 6 to 23 months and 24 to 59 months during a season with a suboptimal match between vaccine and circulating strains. The benefit of the vaccine was measurable among children who were fully vaccinated in comparison with those who were unvaccinated and resulted in an estimated decreased risk for medically attended influenza of 52% and 45%, respectively. Older children, those who were aged 24 to 59 months, benefited from partial vaccination (≥14 days since receipt of 1 of 2 recommended doses), but we determined no VE with partial vaccination among younger children. These results support recommendations for influenza vaccination of children and strengthen the evidence of the vaccine’s ability to reduce substantially the burden of disease among this age group.

ACKNOWLEDGMENTS
We thank Jim Alexander, Sharon Bloom, and Jane Seward for guidance with study design and data interpretation; Connie Knight for administrative support; and Erin Murray and Mona Heaven for assistance with statistical and data management.

REFERENCES


Toward Creating Family-Friendly Work Environments in Pediatrics: Baseline Data From Pediatric Department Chairs and Pediatric Program Directors

Heather A. McPhillips, MD, MPH, Ann E. Burke, MD, Kate Sheppard, MPH, Adam Pallant, MD, F. Bruder Stapleton, MD, Bonita Stanton, MD

Department of Pediatrics, University of Washington School of Medicine, Seattle, Washington; Department of Pediatrics, Wright State University School of Medicine, Dayton, Ohio; Department of Pediatrics, Wayne State University School of Medicine, Detroit, Michigan; Pediatric Residency Program, Brown Medical School, Providence, Rhode Island

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

ABSTRACT

OBJECTIVE. The objective was to determine baseline characteristics of pediatric residency training programs and academic departments in regard to family-friendly work environments as outlined in the Report of the Task Force on Women in Pediatrics.

METHODS. We conducted Web-based anonymous surveys of 147 pediatric department chairs and 203 pediatric program directors. The chair’s questionnaire asked about child care, lactation facilities, family leave policies, work-life balance, and tenure and promotion policies. The program director’s questionnaire asked about family leave, parenting, work-life balance, and perceptions of “family-friendliness.”

RESULTS. The response rate was 52% for program directors and 51% for chairs. Nearly 60% of chairs reported some access to child care or provided assistance locating child care; however, in half of these departments, demand almost always exceeded supply. Lactation facilities were available to breastfeeding faculty in 74% of departments, although only 57% provided access to breast pumps. A total of 78% of chairs and 90% of program directors reported written maternity leave policies with slightly fewer reporting paternity leave policies. The majority (83%) of chairs reported availability of part-time employment, whereas only 27% of program directors offered part-time residency options. Most departments offered some flexibility in promotion and tenure.

CONCLUSIONS. Although progress has been made, change still is needed in many areas in pediatric departments and training programs, including better accessibility to quality child care; improved lactation facilities for breastfeeding mothers; clear, written parental leave policies; and flexible work schedules to accommodate changing demands of family life.
During the past 4 decades, the percentage of women who have graduated from US medical schools has risen from 6.9% in 1966 to 47.1% in 2005. In 2000, the Future of Pediatric Education II Project Task Force published its landmark report, “Organizing Pediatric Education to Meet the Needs of Infants, Children, Adolescents and Young Adults in the 21st Century.” The report’s recommendations covered a wide range of issues regarding pediatric education and the pediatric workforce. Specifically, the report noted a critically low number of pediatric scientists and also noted the increase of women in pediatrics. The authors of Future of Pediatric Education II identified that to accommodate best the changing demographics of pediatrics and to encourage more women to enter academic medicine, strategies would be needed to promote the success of women in fellowship training and academia, including coordinated schedules, fair parental leave policies, quality child care, and flexibility in academic advancement. To address more fully and develop further these strategies, the Federation of Pediatric Organizations (FOPO) established the Task Force on Women in Pediatrics. This task force issued “The Report of the Task Force on Women in Pediatrics” (RTFWP), which made a series of recommendations that aim to improve issues of balancing work and family life, enhance productivity, and foster career advancement and individual fulfillment for women and men in pediatrics. The full report, available at www.fopo.org, is predicated on the conviction that “There are many reasons for addressing issues of family balance in the lives of pediatricians during training and practice, including concerns regarding productivity, career advancement, and individual fulfillment. The most compelling reason derives from the central responsibility of our profession. The commitment of pediatrics to the health and well-being of children and youth should encompass the families of those who choose to pursue careers in pediatrics.”

Recommendations were made for the various levels of career development (from medical school training through senior-level career pediatricians) and involved actions by all 7 of the member organizations that constitute FOPO. Three broad areas of need were identified:

Area 1. Providing and promoting family-friendly environments, including access to varied child care options (e.g., infant, toddler, sick-child, back-up, and after-school care), the provision of lactation facilities, offering mechanisms to aid with child and dependent care expenses, and expanding federal loan forgiveness programs.

Area 2. Flexibility in scheduling and progression through training and career paths, such as part-time training, clinical practice, and research; flexibility in progression through residency curriculum; identifying alternative/flexible career progression; credit for independent projects relating to parenting and child care; moving toward a competence-based evaluation of readiness for certification rather than a required duration of training; and expanding the age requirements for investigator awards to allow for prolonged training or leaves of absence for parenting or elder care.

Area 3. Address issues related to balancing family and work and considerations that are unique to women in the workforce through formal and informal curriculum, mentoring, and career counseling to educate regarding choosing a specialty, defining academic success, negotiating for resources that are necessary for productive academic careers, establishing links to resources to allow or encourage women to pursue research and scholarly activities, and acquiring management skills that are needed for career advancement; and development and dissemination of a recognizable measure or rating system of “family-friendly environment” in medical schools, residencies, subspecialty training, and practice settings as well as data that describe the demographics (e.g., gender, race) of faculty by rank and leadership positions within FOPO.

The Task Force charged all 7 FOPO member organizations to respond to the RTFWP in general, as well as charged individual member organizations with specific recommended actions. Two of the organizations that constitute FOPO, the Association of Medical School Pediatric Department Chairs (AMSPDC) and the Association of Pediatric Program Directors (APPD), initiated their responses through a survey of their membership to ascertain baseline practices that are relevant to the 3 broad categories that were identified in the RTFWP and summarized above. AMSPDC is composed of 147 pediatric department chairs in Canada and the United States. APPD’s membership includes 203 pediatric program directors (PPD) as well as department chairs, assistant program directors, and others who are interested in residency education. This article describes the information that was obtained from these 2 separate surveys and provides recommendations for future actions.

METHODS

AMSPDC initiated its response to the RTFWP with an anonymous Web-based survey of all AMSPDC members (pediatric department chairs) that is designed to describe the current practices, benefits, and policies that are relevant to the recommendations within the RTFWP. Three e-mail messages were sent to the AMSPDC membership in a period of 6 weeks to encourage participation in the survey. The directions to the survey asked the chair to complete the questions to the best of his or her ability. Chairs were encouraged to ask for assistance in respond-
ing to questions to which they were not certain of the correct answer. The questionnaire included 27 major questions that addressed the broad issues of child care, lactation facilities, family leave policies, work-life balance, and tenure and promotion policies.

The APPD independently initiated its response to the RTFWP through a 22-question Web-based survey that addressed family leave, parenting, work-life balance, and perceptions of family-friendliness. All members of the APPD were asked to complete the survey via notification through the periodic APPD electronic Listserv newsletter. Only the responses from self-identified PPDs were used to ensure that each program was counted only once in the results. This research was approved by the institutional review board at Wayne State University.

We used descriptive statistics for most of the data presented. χ² tests were used to compare categorical values. Statistical significance was considered present at P < .05. Stata 8 (Stata Corp, College Station, TX) was used for analysis of the AMSPDC data. SPSS 13.0 (SPSS, Chicago, IL) was used for the APPD data.

RESULTS
Among the 147 AMSPDC members, 75 (51%) returned completed surveys. In the APPD survey, 106 (52%) of the 203 PPDs sent in completed responses.

Area 1: Providing and Promoting Family-Friendly Environments

Infant/Child Care Provisions
Of the 75 department chairs, 59% reported availability of on-site or off-site child care facilities or assistance with locating quality child care for faculty in their department (Table 1). Availability of before- and after-school care, sick-child care, and back-up or emergency day care was more limited. For example, only 13% of department chairs reported that their institution had child care available for a mildly sick child, and only 12% had emergency or back-up child care options. Availability of child care was not associated with department size (P = .7).

Among pediatric departments that offer some form of child care, 50% of chairs responded that demand for child care “always or almost always” exceeded availability, 22% said “occasionally,” and none said “never.” The majority of child care facilities did not reserve spaces for faculty or trainees (75%). Eligibility for enrollment by work/student status was described as follows: 96% allowed enrollment of faculty infant/children; 91% allowed fellows; 89% allowed residents; and 67% permitted medical students. In the APPD survey, 44% of PPDs indicated that their institution offered some “day care opportunities for the children of residents.”

Department chairs reported that few programs offer scholarships or financial subsidization of child care costs (13%). Similar results were provided in the APPD survey; 4% of PPDs responded that their institution subsidizes child care costs for residents, and 2% stated that the institution paid for it.

Almost all infant/child care centers that were available through pediatric departments accepted children aged 6 months and older (94%); however, only 61% accepted them by 2 months of age (Table 2). Typically, child care centers accepted children until they were entering kindergarten.

Among the 44 pediatric departments that provided child care, the majority were governed by the hospital alone (23%) or by the university and hospital (27%). Some departments offered child care through community and other sources (23%).

### TABLE 1 Types of Child Care Provided by 75 Pediatric Departments (AMSPDC Survey)

<table>
<thead>
<tr>
<th>Parameter, n (%)</th>
<th>Yes</th>
<th>No or Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any child care facilities or assistance provided?</td>
<td>44 (59)</td>
<td>31 (41)</td>
</tr>
<tr>
<td>On-site child care</td>
<td>25 (33)</td>
<td>50 (66)</td>
</tr>
<tr>
<td>Off-site child care</td>
<td>18 (24)</td>
<td>57 (76)</td>
</tr>
<tr>
<td>Assistance with locating quality child care</td>
<td>7 (9)</td>
<td>68 (91)</td>
</tr>
<tr>
<td>Before- and/or after-school care available?</td>
<td>14 (19)</td>
<td>61 (81)</td>
</tr>
<tr>
<td>Sick-child care available?</td>
<td>10 (13)</td>
<td>65 (87)</td>
</tr>
<tr>
<td>Back-up or emergency child care available?</td>
<td>9 (12)</td>
<td>66 (88)</td>
</tr>
<tr>
<td>Department size and child care availability</td>
<td>3 2 19 (61)</td>
<td>3 10 (10)</td>
</tr>
<tr>
<td>Minimum age at enrollment in child care, mo</td>
<td>2 19 (61)</td>
<td>3 10 (10)</td>
</tr>
<tr>
<td>≥6 (responses ranged from 18–36)</td>
<td>6 7 (23)</td>
<td>2 (6)</td>
</tr>
</tbody>
</table>

Data may not add up to 100% because of rounding, missing, or “unknown” responses.

*Respondents could choose >1 answer if >1 option applied to their department.

*Respondents could choose >1 answer if >1 option applied to their department.
**Lactation Facilities**

A total of 74% of the AMSPDC respondents reported that their institutions offered lactation facilities, whereas 21% reported no designated space for breast pump use (5% were unsure). Breast pumps were provided by 57% of hospitals/departments for use by faculty and trainees; however, 30% did not provide breast pumps for faculty/trainee use (12% were unsure).

**Family-Oriented Benefits to Parents Who Are Not in a Traditional Husband–Wife Union**

Nearly one third of APPD survey respondents stated that their residency program’s family leave policies and benefits extended to domestic partnerships (32%) or same-gender unions (31%); however, 39% of program directors responded that they did not know whether leave policies or benefits in their program extended to parents in nontraditional unions (not married).

**Area 2: Flexibility in Scheduling and Progression Through Training and Career Paths**

**Leave Provisions**

Of 75 AMSPDC respondents, 78% reported that their department has a written maternity policy and 61% reported that they had a written paternity leave policy. The majority (63%) of chairs responded that their department did not offer additional benefits for parental leave beyond those that are required by the Family Medical Leave Act (Table 3). Paid maternity leave ranged from only accumulated vacation and/or sick time (32%) to >12 weeks (17%). Paid paternity leave ranged from accumulated vacation and/or sick time (46%) to >12 weeks (7%).

Ninety percent of the responding PPDs reported that there is a written maternity leave policy at their training program. PPDs responded that the mean amount of time off that a resident could take without making up time was 3 weeks and the median amount of time reported was 0 weeks (SD: ±3 weeks), with a range of 0 to 12 weeks. On average, PPDs reported that waivers had been requested from the American Board of Pediatrics (ABP) only 1 time (SD: ±2.3 times) in the past 3 years, with a range of 0 to 15. Fifty percent (n = 52) of the responding PPDs reported that they would be in favor of granting some form of academic credit for a “maternity leave elective,” whereas 21% would oppose such an elective. Nearly one third (29%) reported that it would “depend on circumstances.” There was no association between size of the residency program and amount of leave time reported by the program director (Table 4).

**Employment Flexibility**

As shown in Table 3, the majority of department chairs reported that they offer part-time employment for parents (83%) and for faculty who take care of an ill or dependent family member (79%). Other types of employment flexibility were less common. Only 9% of chairs reported that telecommuting jobs were available at their institution, and an additional 21% allowed work that was done at home to count toward daily work hours. The majority (64%) of chairs responded that their department had no official work-from-home options.
although many stated that exceptions were made on a case-by-case basis.

**Part-Time Residencies**

In the APPD survey, 73% of 106 PPDs reported that they did not offer part-time residency options, whereas 27% did. Among those who offered a part-time residency option, the average number of residents who have used that option in the past 10 years was $1.95$ with a range of 0 to 6.

**Promotion and Tenure**

Forty-nine percent of chairs responded that there had been no change in their tenure system in the past 10 years, 39% said that there had been changes, and 12% were unsure. As shown in Table 3, most pediatric departments have some flexibility with the promotion process, either by faculty-initiated ability to stop the “tenure” clock or by longer promotional processes for part-time faculty. In addition, a few chairs reported no firm promotion clock at their institution for clinical faculty. Only 5% of chairs responded that their institution has no flexibility in full-time equivalent or promotion and tenure for issues related to parenting, and 8% stated that there was no flexibility for issues related to the care of an ill or dependent family member.

**Area 3: Issues Related to Balancing Family and Work and Considerations That Are Unique to Women in the Workforce**

**Stress Management**

Most (95%) PPDs believed that it is the PPD’s responsibility to monitor resident well-being. Greater than 90% of responding PPDs acknowledged that programs should be responsible for offering to their trainees guidance on stress reduction, time management, conflict management, and mental health. Fewer PPDs believed that residency programs should be responsible for offering guidance on financial planning (57%) and social activities (50%). Only 45% of PPDs responded that residency programs should offer guidance on child care.

**Balancing Work and Family**

In the AMSPDC survey, 40% of chairs responded that they offer formal instruction on issues of balancing family and work; 7% listed informal mentoring programs or interest groups. However, 52% did not offer formal instruction on balancing family and work.

**Family-Friendly Rating**

Forty-eight percent of PPDs reported that they would be in favor of a rating score with regard to family-friendliness of a residency training program, 23% would not, and 30% were unsure.

**DISCUSSION**

The report issued by the FOPO Task Force on Women in Pediatrics urged that the leading organizations that are concerned with pediatric training and practice seriously consider the implications of the changing workforce. Two of these organizations, AMSPDC and APPD, immediately and independently responded to the report by assessing aspects of the current situation over which they could exert control and change. In addition, both organizations devoted substantial portions of their 2006 annual meetings to this topic, underscoring their recognition of its importance to the future of academic pediatrics and of pediatrics and child health more broadly.

The results of these 2 surveys indicate that change is needed in many areas, including better accessibility to quality child care; improved lactation facilities for breastfeeding mothers; clear, written parental leave policies; and flexible work schedules to accommodate changing demands of family life. In our study, only one third of medical institutions offered on-site child care despite the evidence demonstrating both its positive influence on work performance and importance for both mother and infant of maternal contact, particularly during the first year of life. In many institutions with child care options, demand almost always exceeded availability. Furthermore, child care for a mildly ill child or emergency/back-up child care options rarely were available. Given what is known about maternal stress and mothering as well as stress in employment performance, child care issues should be addressed more aggressively by most, if not all, academic medical complexes.

Supporting women to breastfeed through the first year of their infant’s life is a high priority for pediatricians. Lactation facilities are 1 mechanism to help

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Paid Leave</th>
<th>Unpaid Leave</th>
<th>Paid Leave Only for Time Accrued</th>
<th>Depends on Circumstances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth or adoption of child</td>
<td>57 (54)</td>
<td>31 (30)</td>
<td>36 (34)</td>
<td>8 (8)</td>
</tr>
<tr>
<td>Family medical event</td>
<td>36 (37)</td>
<td>25 (26)</td>
<td>35 (36)</td>
<td>19 (19)</td>
</tr>
<tr>
<td>Death of family member</td>
<td>57 (58)</td>
<td>17 (17)</td>
<td>28 (28)</td>
<td>12 (12)</td>
</tr>
<tr>
<td>Mental health leave</td>
<td>41 (41)</td>
<td>27 (27)</td>
<td>34 (34)</td>
<td>18 (18)</td>
</tr>
<tr>
<td>Personal leave</td>
<td>20 (20)</td>
<td>38 (38)</td>
<td>28 (28)</td>
<td>30 (30)</td>
</tr>
<tr>
<td>Disability or serious illness</td>
<td>57 (55)</td>
<td>25 (24)</td>
<td>32 (31)</td>
<td>11 (11)</td>
</tr>
</tbody>
</table>

Respondents could choose >1 answer if >1 option applied to their program.
support continued nursing by working mothers. The 2005 revision of the American Academy of Pediatrics Policy Statement on Breastfeeding urges pediatricians to "encourage employers to provide appropriate facilities and adequate time in the workplace for breast-pumping." Nonetheless, 1 in 5 pediatric departments in our study currently do not offer lactation facilities to their faculty or trainees.

Given the age of the majority of pediatric residents (70% of whom are women) and junior faculty, childbearing is common during residency, fellowship, and the first stages of academic careers. Uncertainties of parental leave policies and cost of unpaid leave may be enormous sources of stress to parents. Navigating the waters of employment and child rearing, for which there are no written rules, is more complex for residents, fellows, and faculty. To be effective in counseling residents and faculty, academic leaders must be aware of existing policies related to family leave. If no written policies exist, then they should be developed and disseminated. Likewise, in our surveys, both the chairs and the PPDs reported very low levels of child care subsidization; child care costs may be a very real source of anxiety to residents who are parents. The United States differs from other industrialized nations in North America and Europe with regard to both child care subsidies and parent leave policies. Pediatric faculty and residents should encounter positive child care and parental experiences to enhance their effectiveness as advocates for patients and society. The benefits of subsidized child care should be determined independently for both faculty and residents or other trainees and compared with the potential negative impact of viewing this as an unequal or unfair benefit by employees who do not need access to child care. Given the financial vulnerability of trainees compared with most faculty, a need-based system that is open to all employees might be a possible solution.

Family, accommodations within the “tenure clock” vary widely between institutions, ranging from a few that offer no flexibility for parents, to those that have abolished the tenure clock altogether. Although the majority of institutions offer part-time employment options for faculty, not all do so and somewhat fewer seem to offer it for other dependent care. In addition, it is unclear how part-time options may effect faculty on who are different tracks (eg, research tracks, clinician tracks). Much has been written regarding tenure and part-time work and regarding parenting and tenure in the past decades. The huge range in current practices suggests that more work is needed to address this substantial problem as increasing numbers of women enter medicine.

Part-time residency training is offered in one quarter of the programs in our study, but, on average, <1 resident per program used this part-time option. This finding is consistent with other published data. Data that were obtained from national surveys that were conducted by the Accreditation Council for Graduate Medical Education and from the Graduate Medical Education Track/Fellowship and Residency Electronic Interactive Database system are similar to findings in our study. The Accreditation Council for Graduate Medical Education reported that in 2004–2005, there were 7795 total pediatric residents and 15 (<0.2%) were in part-time residency programs. In the Graduate Medical Education Track/ Fellowship and Residency Electronic Interactive Database data for 2004–2005, 44 (22%) of 204 pediatric residency programs stated that they offer part-time or shared positions, compared with 27% of respondents in our survey, conducted in early 2006. It is unclear why such a seemingly family-friendly option is taken advantage of so infrequently. A variety of reasons might make the option less appealing, including increased time in training, financial issues and/or guilt of perceived increased workload on colleagues. Several mechanisms for part-time residency training or “shared” residencies have been described. Research into the effects, implementation, and barriers of part-time training on residency programs, individual residents, and institutions is needed. Furthermore, the resident who is a parent and does not want or is unable to work/train on a part-time basis may have a different set of stressors to contend with during training, including guilt of not being involved on a daily basis with child rearing. Efforts should be made to determine how best to support full-time resident parents to achieve both the maximum benefit of residency education and adequate time with their family.

The APPD survey results also highlight the variability among programs in the interpretation and adherence to the ABP requirements that a resident complete 33 months of residency. One month of absence is allowed each year by the ABP, including time spent on vacation, for parental leave, or for illness. For example a resident who starts on July 1, 2006, and finishes internship on June 30, 2007, would have completed 11 months if he or she took 3 weeks of vacation and 1 week during the winter holidays. A resident who is always present on his or her assigned rotations, except for designated vacations, typically completes 33 months of training during 3 years of residency. Any other time absent should be made up or formally waived by the ABP. Our results show that some PPDs adhere strictly to this rule; others do not. This may be because “block” rotations do not align perfectly with calendar months in many residency programs, allowing perhaps for different interpretations of rules for leave (eg, 1 month may or may not be counted as equivalent to 4 weeks).

**Limitations**

There are several limitations to the present study, including the 51% to 52% response rate to both surveys. Because this survey was anonymous, it is not possible to compare characteristics of pediatric departments or res-
idency programs between respondents and nonrespondents. It therefore is possible that our data are not wholly representative of characteristics of family-friendly departments and residency programs because of response bias. Neither survey was constructed from previous qualitative research. Likewise, we did not conduct psychometric testing of the questions asked.

Implications

These data raise important questions regarding parenting for pediatric residents and faculty. Our data demonstrate a wide diversity in the presence of institutional policies and programs that focus on creating family-friendly environments for residents and faculty. Even when present, faculty and residents frequently do not take advantage of these unique opportunities. Understanding the feelings and attitudes toward the dual roles of parent and pediatrician would allow the pediatric community to understand more clearly family-friendly issues that involve child care and career progression. AMSPDC and APPD should take the lead in determining “best practices” for family-friendly academic departments and residency programs. To begin with, additional exploration of barriers, perceived and real, to developing family-friendly policies is needed.

Our surveys show the need for more accessible and flexible child care that not only allows but also encourages parent–child interaction and nurturing, universal accessibility to lactation facilities for breastfeeding mothers at work, and greater flexibility in training and academic career progression. Implementation of these initiatives does not require more research. Although some of the changes must have the commitment from the entire academic institution, change must begin somewhere. Where better to begin than within the departments and programs that are most concerned about children and families?

REFERENCES

Incidence, Prognosis, and Risk Factors for Fatigue and Chronic Fatigue Syndrome in Adolescents: A Prospective Community Study

Katharine A. Rimes, DPhil*, Robert Goodman, PhDb, Matthew Hotopf, PhD*, Simon Wessely, MDa, Howard Meltzer, PhDc, Trudie Chalder, PhDa

King’s College London, Institute of Psychiatry, *Section of General Hospital Psychiatry and bDepartment of Child and Adolescent Psychiatry, London, United Kingdom; cAcademic Department of Psychiatry, Brandon Mental Health Unit, Leicester General Hospital, Leicester, 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 incidence, prevalence, risk factors, and prognosis of fatigue, chronic fatigue, and chronic fatigue syndrome in 11- to 15-year-olds.

METHODS. A random general population sample (n = 842) of British adolescents and their parents were assessed at baseline and 4 to 6 months later. The main outcomes were fatigue, chronic fatigue, and chronic fatigue syndrome, operationally defined.

RESULTS. The incidence over 4 to 6 months was 30.3% for fatigue, 1.1% for chronic fatigue, and 0.5% for chronic fatigue syndrome. The point prevalence was 34.1% and 38.1% for fatigue, 0.4% and 1.1% for chronic fatigue, and 0.1% and 0.5% for chronic fatigue syndrome at time 1 and time 2, respectively. Of participants who were fatigued at time 1, 53% remained fatigued at time 2. The 3 cases of chronic fatigue and 1 case of chronic fatigue syndrome at time 1 had recovered by time 2. Higher risk for development of chronic fatigue at time 2 was associated with time 1 anxiety or depression, conduct disorder, and maternal distress; in multivariate analysis, baseline anxiety or depression remained a significant predictor of chronic fatigue. Increased risk for development of fatigue at time 2 was associated with time 1 anxiety or depression, conduct disorder, and older age; in multivariate analyses, these factors and female gender all were significant predictors of fatigue.

CONCLUSIONS. The incidence rates for chronic fatigue and chronic fatigue syndrome in this adolescent sample were relatively high, but the prognosis for these conditions was good. This prospective study provides evidence for an association between emotional/behavioral problems and subsequent onset of fatigue/chronic fatigue.
CHRONIC FATIGUE SYNDROME (CFS) is characterized by disabling physical and mental fatigue that lasts at least 6 months, with a number of accompanying symptoms. The term “chronic fatigue” (CF) has no widely agreed-on definition but is used variously to describe excessive, persistent fatigue that is associated with less impairment than CFS; fatigue that has fewer accompanying symptoms; or fatigue that has lasted <6 months. There now is an extensive empirical literature on many aspects of CF and CFS, but nearly all relates to adults.

Estimates of prevalence in adults vary widely, ranging from ~1.8% to 9% for CF and ~0.075% to 0.42% for CFS in community studies, depending on the setting, methods, and definitions used. Using data from the British Child and Adolescent Mental Health Survey 1999 conducted by the Office for National Statistics, we reported that in 5- to 15-year-olds, CF was present in 0.57% of the sample and CFS according to Centers for Disease Control and Prevention (CDC) criteria was present in 0.19%. A previous incidence study in adults reported an annual incidence of 50 per 1000 per year for CF and 3.7 per 1000 per year for CFS. Incidence rates for fatigue, CF, and CFS in adolescents are not known.

The present study involves the longitudinal follow-up of the previously used cohort to provide data on the incidence, prognosis, and risk factors for fatigue, CF, and CFS when assessed 4 to 6 months after the original study in a subgroup of those participants.

METHODS
Design and Participants
This study was part of a larger Office for National Statistics study of mental health in children who were aged 5 to 15 years. The sample frame consisted of families who lived in private households in England, Scotland, and Wales. The sample was taken from child benefit records. The original sampling frame excluded cases for which “administrative action” such as death or change of address was indicated. The sampling frame was stratified by Regional Health Authority and within that by sociodemographic groupings. Postal sectors were randomly selected within that frame with a probability proportional to the size of the sector. In each of the 475 postal sectors, 30 children were randomly selected and targeted. The Child Benefit Centre sent information to parents about the survey, giving them the opportunity to opt out. Those who agreed to participate were interviewed at home.

Of the 14,250 families contacted, 931 (6.5%) opted out by calling the Child Benefit Register and 790 (5.5%) addresses were found to be incorrect. This left a sample of 12,529 eligible children, and 10,438 of these were interviewed. Failed interviews were attributed to either non-contact (2%) or refusal (15%). Of those interviewed, 4,240 were 11- to 15-year-olds and were included in the original epidemiologic study that was reported in the previous article. From this first time point, 1,096 adolescents were selected to be reassessed 4 to 6 months later for the present study. Follow-ups were restricted to participants who had been assessed in the early stages of the study because of budgetary constraints and because this permitted a 4- to 6-month follow-up before the end of the school year. Follow-up questionnaires were returned by 77% (n = 842) of this sample.

Measures and Interviews
At time 1, the Development and Well-being Assessment was administered by nonclinical, trained interviewers. Classification according to the International Classification of Disease, Tenth Revision, Classification of Mental and Behavioral Disorders with strict impairment criteria was ascertained by a combination of interviews and rating techniques that were designed to generate psychiatric diagnoses on this age group. The Development and Well-being Assessment has been shown to discriminate well between community and clinic samples in rates of diagnosed disorders. The adolescents were asked whether they had been feeling more tired and worn out than usual. When they answered affirmatively, the interviewer asked supplementary and open-ended questions related to duration, effect of fatigue on various aspects of their life, and number and severity of symptoms, including questions to assess CDC criteria for CFS (Appendix E, questions C3D1–C3D199). These supplementary questions included 4 items concerning impairment: “Has feeling tired and worn-out... interfered with (1) how well you get on with the rest of your family, (2) making and keeping friends, (3) learning or class work, (4) playing, hobbies, sports or other leisure activities?” For each of the 4 domains, the participant could choose from among the responses “not at all,” “only a little,” “quite a bit,” and “a great deal.” The responses “quite a bit” and “a great deal” were coded as significant impairment and received scores of 1 and 2, respectively. The total impairment score was the sum of these 4 items and therefore ranged from 0 to 8. Mothers completed the 12-item General Health Questionnaire (GHQ) to assess maternal psychological well-being. The British Picture Vocabulary Scale was used to assess the adolescent’s IQ. Follow-up was conducted by mail and included the same questions that were related to fatigue and asked at time 1.

Outcomes
The adolescent was classified as fatigued when he or she answered “yes” to the question, “Over the last month, have you been feeling much more tired and worn out than usual?” CF was defined operationally as severe fatigue that was of at least 6 months’ duration, was not helped by rest, and was associated with significant im-
impairment in at least 1 domain (see previous section). CFS was defined according to the CDC criteria.  

### Analysis

Statistical analysis was conducted without adjustment for stratification and clustering because previous experience has shown that prevalence figures are minimally affected by these processes in this sample. Logistic regression (also with no adjustment for stratification and clustering) was used to calculate odds ratios between independent variables at time 1 and dependent fatigue variables at time 2. A final multivariate model was used to determine which factors were associated independently with fatigue status at time 2, in which all variables that were shown to be associated significantly \((P < .05)\) were entered simultaneously, together with gender and age, because these were known confounders.

### RESULTS

#### Characteristics of Participants

Characteristics of participants who took part in the study at both times 1 and 2 were compared with those for the rest of the 11- to 15-year-old participants at time 1. T tests indicated that time 2 participants had a significantly higher IQ score on the British Picture Vocabulary Scale than those who only took part in the first part of the study (mean: 102.02 vs 98.72; \(t_{4174} = 4.7; P < .0005\)). Time 2 participants were more likely to come from a higher social class (58.1% vs 53.4%; \(t_9273 = 5.7, P < .017\)), to have a mother educated to at least “A” levels (advanced level exams) or equivalent (34.4% vs 30.6%; \(t_9273 = 4.4, P < .037\)), and to have a mother who had a partner (82.3% vs 75.6%; \(t_9273 = 17.1, P < .0005\)) but were less likely to report nonwhite ethnicity (7.0% vs 9.5%; \(t_9273 = 5.0, P = .026\)) or to have a psychiatric diagnosis at time 1 (7.5% vs 10.5%; \(t_9273 = 6.8, P = .009\)) than the other participants. There were no significant differences between the 2 groups with regard to age (at 1 day \(t_{4238} = 1.5\), gender (\(t_{4152} = 0.02\)), maternal GHQ score (\(t_{4152} = 0.5\)), or whether the participant reported being tired at time 1 (\(t_{1} = 2.2\)).

#### Incidence of Fatigue, CF, and CFS

There were 166 new cases of fatigue at time 2 (30.3%; 95% confidence interval [CI]: 26.4–34.2) and 9 new cases of CF (1.1%; 95% CI: 0.04–1.8) and 4 new cases of CFS (0.5%; 95% CI: 0.01–0.9) in the 4- to 6-month period (Table 1).

#### Point Prevalence of Fatigue, CF, and CFS

The point prevalence of fatigue was 34.1% (95% CI: 30.9–37.3) at time 1 and 38.1% (95% CI: 34.8–41.5) at time 2. For CF, the point prevalence was 0.4% (95% CI: 0.0–0.8) at time 1 and 1.1% (95% CI: 0.04–1.8) at time 2. For CFS, the point prevalence was 0.1% (95% CI: 0.0–0.4) for time 1 and 0.5% (95% CI: 0.01–0.9) for time 2 (Table 1).

#### Prognosis of Fatigue, CF, and CFS

Of the 283 participants who reported that they were tired at time 1, 151 (53.3%; 95% CI: 47.5–59.2) were still tired at time 2. All 3 participants with CF and the 1 participant with CFS at time 1 had recovered by time 2 (Table 1).

#### Impairment That Was Associated With Fatigue, CF, and CFS

Of the participants with fatigue at time 1, 22.6% (95% CI: 17.7–27.5) reported significant impairment in at least 9.5%; \(X_1 = 5.0, P = .026\) or to have a psychiatric diagnosis at time 1 (7.5% vs 10.5%; \(X_1 = 6.8, P = .009\)) than the other participants. There were no significant differences between the 2 groups with regard to age (at 1 day \(X_{4238} = 1.5\), gender (\(X_{4152} = 0.02\)), maternal GHQ score (\(X_{4152} = 0.5\)), or whether the participant reported being tired at time 1 (\(X_{1} = 2.2\)).

### Table 1

<table>
<thead>
<tr>
<th>Total Classifications at Time 1</th>
<th>Classification at Time 2 Divided According to Classification at Time 1</th>
<th>Total Classifications at Time 2 (4–6 mo Later)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue</td>
<td>283 (34.1%) fatigued</td>
<td>151 (53.4%) fatigued</td>
</tr>
<tr>
<td></td>
<td>548 (65.9%) not fatigued</td>
<td>132 (46.6%) not fatigued</td>
</tr>
<tr>
<td>CFb</td>
<td>3 (0.4%) CF</td>
<td>0 (0%) CF</td>
</tr>
<tr>
<td></td>
<td>839 (99.6%) no CF</td>
<td>9 (1.1%) CF</td>
</tr>
<tr>
<td>CFSc</td>
<td>1 (0.1%) CFS</td>
<td>0 (0%) CFS</td>
</tr>
<tr>
<td></td>
<td>841 (99.9%) no CFS</td>
<td>1 (100%) No CFS</td>
</tr>
<tr>
<td></td>
<td>a New cases.</td>
<td>b Defined as severe fatigue of 6 months’ duration with nonrestorative rest and functional impairment.</td>
</tr>
<tr>
<td></td>
<td>c According to CDC criteria.</td>
<td></td>
</tr>
</tbody>
</table>

---

\(^a\) New cases.  
\(^b\) Defined as severe fatigue of 6 months’ duration with nonrestorative rest and functional impairment.  
\(^c\) According to CDC criteria.
Time 1 Factors That Were Associated With New Fatigue at Time 2
To investigate which factors at time 1 were associated with new-onset fatigue at time 2, hypothesized risk factors were entered into a logistic regression model (see Table 2 for details). In univariate analyses, older age, anxiety or depressive disorder, and conduct disorder at time 1 were associated significantly with fatigue at time 2. When these variables were entered into a multivariate model together with gender, the results indicated that age, female gender, conduct disorder, and anxiety or depression were significant independent predictors of fatigue.

Time 1 Factors That Were Associated With Persistence of Fatigue at Time 2
Logistic regression analyses were performed using participants who were fatigued at time 1 to investigate which factors were associated with either persistent fatigue at time 2 or recovery (see Table 2 for details). Individuals who still reported being much more tired or worn out than usual in the past month at time 2 were regarded as having persistent fatigue, whereas those who did not report fatigue at time 2 were regarded as having recovered. Time 1 factors that were associated with remaining fatigued were older age, female gender, higher IQ, higher maternal GHQ score, and the mother’s being educated to at least A levels or equivalent. When these variables were entered together in a multivariate analysis, older age, female gender, higher IQ, and maternal GHQ score were significant independent predictors of persistent fatigue.

Characteristics of Individuals With Recent-Onset CFS
There were only 4 new cases of CFS at time 2, so the numbers were too small to analyze risk factors statistically. Three of the 4 new cases of CFS were in girls. Three of the 4 had at least 1 psychiatric diagnosis at baseline. Three of them had reported being “much more tired and worn out than usual over the last month” at time 1. Two of the participants had frequent headaches at time 1; 1 of them also had sleep problems and postexertional malaise at this point.

### TABLE 2  Associations Between Factors at Time 1 and New-Onset Fatigue, Persistent Fatigue, and New-Onset CF at Time 2

<table>
<thead>
<tr>
<th>Factors Assessed at Time 1</th>
<th>Participants With Fatigue at Time 1: New Cases of Fatigue at Time 2 ( (n/N = 166/548) )</th>
<th>Participants With Fatigue at Time 1: Prognosis of Fatigue at Time 2 ( (n/N = 151/283) )</th>
<th>Participants With Chronic, Severe Fatigue With Disability at Time 2 ( ^a (n/N = 9/839) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.30 (1.14–1.49) (^b)</td>
<td>1.34 (1.16–1.54) (^b)</td>
<td>1.35 (0.82–2.22) (^a)</td>
</tr>
<tr>
<td>Female gender</td>
<td>1.43 (0.99–2.07)</td>
<td>1.77 (1.20–2.61)</td>
<td>0.82 (0.22–3.07)</td>
</tr>
<tr>
<td>Social class (I, II, III versus IIIM, IV, V)</td>
<td>1.04 (0.72–1.51)</td>
<td>0.74 (0.46–1.21)</td>
<td>0.46 (0.09–2.29)</td>
</tr>
<tr>
<td>Ethnicity (white versus nonwhite)</td>
<td>1.16 (0.57–2.38)</td>
<td>1.98 (0.73–5.36)</td>
<td>1.66 (0.21–13.53)</td>
</tr>
<tr>
<td>IQ</td>
<td>1.00 (0.99–1.01)</td>
<td>1.02 (1.01–1.04)</td>
<td>0.99 (0.95–1.03)</td>
</tr>
<tr>
<td>Any anxiety or depressive disorder</td>
<td>4.88 (1.80–13.2)</td>
<td>4.33 (1.49–12.6)</td>
<td>30.2 (7.75–117) (^b)</td>
</tr>
<tr>
<td>Any conduct disorder</td>
<td>3.16 (1.36–7.36)</td>
<td>3.01 (1.18–7.68)</td>
<td>2.45 (0.48–103) (^b)</td>
</tr>
<tr>
<td>Hyperkinetic disorder</td>
<td>0.57 (0.06–5.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tired at time 1</td>
<td>1.00 (0.96–1.04)</td>
<td>1.02 (1.01–1.04)</td>
<td>0.99 (0.95–1.03)</td>
</tr>
<tr>
<td>Mother’s GHQ score</td>
<td>1.05 (1.00–1.12)</td>
<td>1.09 (1.00–1.17)</td>
<td>1.18 (1.01–1.38)</td>
</tr>
<tr>
<td>Mother educated “A” level or above</td>
<td>1.05 (0.71–1.55)</td>
<td>1.80 (1.11–2.92)</td>
<td>2.40 (0.64–9.00)</td>
</tr>
<tr>
<td>Mother: single (versus partner)</td>
<td>1.53 (0.97–2.41)</td>
<td>1.82 (0.96–3.46)</td>
<td>3.78 (1.00–14.3)</td>
</tr>
</tbody>
</table>

\(^a\) Severe fatigue of at least 6 months’ duration with nonrestorative rest and functional impairment.

\(^b\) \( P < 0.005 \)

\(^c\) \( P < 0.005 \)

\(^d\) \( P < 0.05 \)

\(^e\) Not possible to analyze because cell sizes were too small.
DISCUSSION
We have already shown that fatigue is common in adolescents but CF and CFS less so, a finding that since has been replicated elsewhere. In this article, we present the first incidence data, the first prospective epidemiologic data on prognosis, and the first prospective data on preillness risk factors for fatigue, CF, and CFS in adolescents. The use of a community sample ensures that the findings are less open to the selection and referral biases that may operate in clinical samples. The prospective design allows identification of predictor/risk factors and can help in determination of the sequence of causality. This is particularly important for factors such as anxiety or depression, which may occur as a consequence of conditions such as CF or CFS. However, using a prospective design shows that these and other characteristics clearly are vulnerability factors for fatigue conditions.

The incidence rates for CF and CFS were surprisingly high (1.1% for CF and 0.5% for CFS during a 4- to 6-month period), given the relatively low prevalence. Prevalence is a function of incidence and prognosis, and we found that the high incidence was balanced by high remission rates 4 to 6 months later.

Having an anxiety or depressive disorder at time 1 was strongly associated with new onset of fatigue or CF at time 2. Conduct disorder was associated with new-onset fatigue and CF in univariate analyses and remained a significant predictor of new fatigue (but not CF) in multivariate analysis. Three of the 4 adolescents who developed CFS at time 2 had had at least 1 psychiatric diagnosis at time 1. Maternal distress at time 1 was associated with the persistence of fatigue and new-onset CF at time 2. Older age and female gender also were associated with an increased risk for new-onset fatigue and the persistence of fatigue between times 1 and 2. Higher IQ and having a mother who was educated to A levels or beyond were associated with the persistence of fatigue.

Limitations
One limitation is that the participants from time 1 who also completed the questionnaire at time 2 were different from the rest of the time 1 sample in several ways that reflected higher socioeconomic status and better mental health. These differences might influence the results. For example, a strong predictor of fatigue or CF at time 2 was having a psychiatric diagnosis at time 1, and because people were less likely to take part at time 2 when they had a psychiatric diagnosis, it is possible that the incidence and prevalence figures at time 2 are an underestimation of those that would have been found if all of the sample had taken part at follow-up. Another limitation is that the small number of participants with CF in this sample means that we cannot draw firm conclusions from these analyses regarding the factors that were associated with the onset of this condition. (Similarly, the numbers of participants with CFS were too small to undertake statistical analysis to examine factors that were associated with onset.) However, 2 of the factors that were identified as predictors of new-onset of CF—conduct disorder and anxiety/depression—also were identified as being associated with new cases of fatigue that were less chronic or disabling, for which the numbers of cases were much larger. An additional limitation is that the diagnoses at time 2 were made on the basis of questionnaire rather than clinical interview data; it was not possible for logistic reasons to repeat the psychiatric interview at time 2. At neither time point were the adolescents physically examined or investigated to exclude other possible diagnoses. However, the number of cases of fatigue that were caused by defined biomedical diagnoses in this age group is likely to be very small.

Comparison With Others Studies
This is the first study, to our knowledge, that examined the incidence of CF and CFS in adolescents. The incidence was 11 per 1000 for CF and 5 per 1000 for CFS for the 4- to 6-month study period, findings that are broadly comparable to the incidence figures that were reported in an adult community sample (50 per 1000 per year for CF and 3.7 per 1000 per year for CFS), despite the use of different definitions for CF and CFS in that study. The incidence rates for CF and CFS in the present study were higher than those previously reported for a number of other disorders in adolescents, such as asthma (estimated annual incidence of 2.1 per 1000 per year), type 1 diabetes (0.16–0.18 per 1000 per year), and anxiety disorders (6.1 per 1000 per year). However, the CF and CFS incidence figures are lower than those for depression in adolescents (33 per 1000 per year).

The prevalence of CF in the adolescents in this sample was slightly lower than that reported for adults in previous community studies and was lower than in a study of children and adolescents in the Netherlands, where 9.6% of girls and 2.3% of boys had severe and chronic fatigue. The latter study used a different definition and assessment of severe fatigue, which, for example, did not require nonrestorative rest or functional impairment, was of 3 months’ duration rather than 6, and included motivational problems. The present study finds higher CFS prevalence rates (0.1% and 0.5%) than the estimated 0.034% in an American adolescent community study, although the figures are within the ranges reported in adult community research. The finding that the adolescents with CF or CFS at time 1 all had recovered by time 2 is consistent with previous tertiary care research indicating that adolescents with CFS have better outcomes than adults.

No previous prospective studies have examined variables that predict the onset of fatigue, CF, and CFS in young people. The demonstration of a relationship be-
tween psychiatric disorder and subsequent fatigue or CF in this study is consistent with previous cross-sectional and retrospective research into fatigue and CFS in adolescents. For example, one study found that three-quarters of young people (aged 10–18) had CFS and attended a specialist clinic had had a psychiatric disorder in the previous year. In a prospective primary care study in adults, participants with depression had an increased risk for a new episode of unexplained fatigue at 12-month follow-up. An interesting finding in the present study was the association between conduct disorder and CF, a relationship that was not demonstrated previously.

The findings that indicated associations between maternal distress and subsequent new CF and persistent fatigue and between older age and female gender and new-onset fatigue are consistent with previous cross-sectional findings in adolescents. Unlike a prospective study of CFS in adulthood, this study did not find an association between social class and the onset of CFS in adolescence. Prospective research into adult-onset CFS has identified other risk factors, including lower childhood physical fitness, long-standing physical or mental illness in childhood that significantly affected home or school life, certain viruses (e.g., mononucleosis), and prolonged convalescence and particular illness beliefs after a virus.

**Implications of the Findings**

The nature of the relationship between psychological problems and subsequent fatigue states requires additional investigation. Possible explanations for the association include overlapping criteria in the definitions of CFS and some psychiatric conditions, psychological problems playing a contributory role in the development of fatigue syndromes, or psychological conditions and fatigue syndromes sharing common risk factors that account for the association between them. One implication of these results is that when health professionals consider the various factors that may be contributing to a young person’s fatigue, of any severity, they should assess for emotional/behavioral difficulties.

Although this study suggests that the prognosis of CF and CFS in adolescents who are aged 11 to 15 may be relatively good, even a few months’ school absence may have profound effects on an adolescent’s education and peer relationships. Furthermore, evidence from tertiary care suggests that a substantial minority of individuals remain disabled for long periods of time. Therefore, effective treatments are needed to help reduce the length of the illness and improve outcomes. Results from a randomized, controlled trial indicate that cognitive behavior therapy can be effective for adolescents who have CFS and are referred from a pediatrics clinic. This intervention involves components that are tailored to the specific needs of the child, such as the establishment of a baseline level of activity and rest (to prevent bursts of overactivity and subsequent severe fatigue), followed by a gradual, structured increase in physical activity and cognitive work to address unhelpful thinking patterns that might be contributing to functional impairment or emotional problems. There is little evidence regarding primary care interventions for adolescents with CF or CFS, but evidence from this study suggests that emotional factors may need to be addressed in this setting too.

**CONCLUSIONS**

This is the first prospective community survey of the incidence, prognosis, and risk factors for CF and CFS in adolescents, thus overcoming the problems that are associated with cross-sectional studies that are conducted in specialist settings. The observed incidence rates for CF and CFS in adolescents were relatively high. Anxiety or depressive disorder was associated with an increased risk for fatigue or CF 4 to 6 months later. Assessment of new cases of fatigue therefore should include a review of any possible psychological problems. Future research needs to investigate the nature of the relationship between the identified risk factors and the development of CF or CFS and identify other risk factors. Although the prognosis for CF and CFS in this sample was good, these conditions can be extremely disabling, and effective treatments for young people in primary care settings are needed.

**ACKNOWLEDGMENTS**

We thank all of the families who kindly agreed to take part.

**REFERENCES**


Early Childhood Gender Differences in Anterior and Posterior Cerebral Blood Flow Velocity and Autoregulation

Nuj Tontisirin, MDa, Saipin L. Muangman, MDa, Pilar Suz, MDa, Catherine Pihoker, MDb, Dana Fisk, RNb, Anne Moore, RVT4, Arthur M. Lam, MDc,d, Monica S. Vavilala, MDa,b,c,d

Departments of aAnesthesiology, bPediatrics, and cNeurological Surgery, University of Washington, Seattle, Washington; dCerebrovascular Laboratory, Harborview Medical Center, Seattle, Washington

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

ABSTRACT

OBJECTIVE. We aimed to describe gender differences in blood flow velocity and autoregulation of the anterior and posterior cerebral circulations in prepubertal children.

METHODS. A prospective observational cohort study was performed at Harborview Medical Center’s Cerebrovascular Laboratory after institutional review board approval, consent, and assent procedures. Children underwent measurement of middle cerebral and basilar artery flow velocities and cerebral autoregulation testing of the middle cerebral and basilar arteries. Cerebral autoregulation was quantified using the autoregulatory index, and estimated cerebrovascular resistance was calculated. Autoregulatory index <0.4 reflects impaired cerebral autoregulation. Data are presented as mean ± SD. Patients were healthy 4- to 8-year-old children.

RESULTS. Forty-eight children (24 boys and 24 girls) 4 to 8 years of age (mean: 6 ± 2 years) were enrolled. Middle cerebral artery flow velocity was higher than basilar artery flow velocity (96 ± 13 vs 65 ± 11 cm/s). Girls had higher middle cerebral artery flow velocity (99 ± 11 vs 91 ± 13 cm/s) and basilar artery flow velocity (70 ± 10 vs 61 ± 9 cm/s) than boys. Cerebral autoregulation was intact in all children. There was no gender difference in autoregulation between the middle cerebral artery (boys: 0.97 ± 0.07; girls: 0.94 ± 0.11) or basilar artery (boys: 0.94 ± 0.13; girls: 0.94 ± 0.11).

CONCLUSIONS. Similar to older children and adults, girls between 4 and 8 years of age had higher middle cerebral and basilar artery flow velocity than age-matched boys. This difference may reflect inherent differences in cerebral metabolic rate and/or estimated cerebrovascular resistance between the genders.
In children, cerebral blood flow (CBFV) normally changes with age. It is low during infancy and increases during childhood before decreasing to adult levels during adolescence. Although normative age-related differences in CBFV are recognized in children, little is known about gender differences in either CBFV or cerebral autoregulation during early childhood.

Adult data indicate that middle cerebral artery flow velocity (Vmca) is higher in women than in age-matched men. In a recent pediatric study of 26 healthy 10- to 16-year-old pubertal adolescents, the investigators reported higher Vmca and basilar artery (BA) flow velocity (Vbas) in girls than in boys. In this study, girls had greater autoregulation of the BA compared with boys, whereas boys had greater autoregulation of the middle cerebral artery (MCA). Although the sample size in this study was small, and the significance of the reported differences in autoregulation remains unclear, these findings suggested that the observed gender-related cerebrovascular differences in adults are present during childhood and might be because of hormonal influences. However, whereas the authors used Tanner staging to document puberty, the age range of the children spanned 6 years, and no data from prepubertal children were included, thereby limiting the ability of the authors to define the age at which these potentially hormonally mediated changes occurred. Therefore, the purpose of the present study was to describe gender differences in anterior and posterior CBFV and autoregulation in prepubertal children during early childhood.

PATIENTS AND METHODS

Study Participants and Setting
This study was approved by the University of Washington’s institutional review board. Healthy American Society of Anesthesiologists category I prepubertal children 4 to 8 years old were eligible. Children with a history of seizures, hypertension, syncope, dysautonomia, or neurologic/cardiac disorders were excluded. Children were recruited from well-child visits in the general pediatrics clinic at Harborview Medical Center. Informed consent for participation was obtained from the parent or guardian, and age-appropriate assent was obtained from each child. Physiologic testing was conducted at Harborview Medical Center’s cerebrovascular laboratory.

Study Design and Protocol
Each subject was positioned on a bed with the head and back adjustable for elevation. An appropriately sized noninvasive blood pressure cuff was placed on 1 arm. Transcranial Doppler (TCD) ultrasonography (Multidop X; DWL Corp, Sipplingen, Germany) was used to measure flow velocities in the middle cerebral (Vmca) and basilar (Vbas) arteries (Figs 1 and 2). A hand-held 2-mHz ultrasound probe was used to insonate the desired vessel and positioned for sufficient time to achieve steady-state measurement. Previously established age-appropriate depths were used to insonate both the MCAs and BAs. During steady-state conditions, Vmca, Vbas, mean arterial pressure (MAP), heart rate, and respiratory rate were recorded in each position.

To examine cerebral autoregulation, change in position proceeded from the supine to sitting upright (90°) position. Five minutes was allowed between position changes before data collection. For the sitting upright position, the vertical distance between the noninvasive blood pressure cuff and the external auditory meatus was used to calculate the estimated MAP (MAPe) at the Circle of Willis. Because MAP decreases by 1 mm Hg for every 1.36-cm increase in vertical height, the change in height from supine to sitting upright was divided by 1.36 cm to calculate the estimated MAPe in the sitting upright position. eCVR (estimated cerebrovascular resistance [CVR]) was calculated as the MAPe divided by Vmca or Vbas.

Briefly, cerebral autoregulation was quantified using the autoregulatory index (ARI) where ARI = Δ eCVR/Δ MAPe; eCVR is calculated as the ratio of MAP to Vmca or Vbas, as appropriate. Two ARIs were calculated for each subject: 1 for the MCA (ARImca) and 1 for the BA (ARIbas).

Sample Size Calculation and Statistical Analysis
Sample size calculation was based on previously published ARI data. We considered a 30% difference in mean ARI to be significant. Assuming α is .05 and β is .8, power analysis indicated that we needed 12 subjects in each group. The normal mean ARI in children without neurologic disease under general anesthesia is 0.75 to 0.90, and a mean ARI ≥0.4 reflects intact cerebral au-
to regulate. CBFV (Vmca and Vbas) and ARI were analyzed by gender. Two-factor analysis of variance was used to compare ARImca versus ARIbas and Vmca versus Vbas between boys and girls. All of the data are presented as mean ± SD.

RESULTS

Forty-eight children (24 boys and 24 girls) between the ages of 4 and 8 years (boys: mean: 6 ± 2 years; girls: mean: 6 ± 2 years; P = .85) were enrolled. There were no adverse events during testing, and all of the children completed the study procedure successfully. There was no overall difference in either blood pressure or respiratory rate during physiologic testing (Table 1).

For both boys and girls, Vmca was higher than Vbas (Vmca: 96 ± 13 cm/s vs Vbas 65 ± 11 cm/s; P = .002; Table 2). In both the supine and upright positions, girls had higher Vmca and Vbas than boys (Table 2). Girls had a lower CVRe of BA than boys (Tables 2 and 3), but there was no gender difference in MCA CVRe.

All of the study participants had intact cerebral autoregulation of the MCAs and BAs. There was no difference in the upright MAPe or percentage drop of MAPe between boys and girls (Table 1). There was no difference in either ARImca or ARIbas between boys and girls (Table 4).

DISCUSSION

The results of this study show that prepubertal 4- to 8-year-old girls have higher Vmca and Vbas than age-matched boys. This gender-related difference is similar to that described previously in older pubertal children (Table 5). However, unlike in older children, we did not find any gender difference in cerebral autoregulation. This is the first study to document that the observed adult gender differences in CBFV are present during early childhood.

TCD ultrasonography is a noninvasive and bedside tool used to estimate CBF and measure CBFV. Clinically, TCD is used to determine CO2 reactivity and cerebral autoregulation, as well as to diagnose stroke, postoperative emboli, and cerebral vasospasm. It is used to prevent stroke in sickle cell disease. Referent age-related Vmca and Vbas data in children with and without disease using TCD ultrasonography are consistent with age-related CBF changes documented by other neuroimaging modalities, such as single photon emission computed tomography and positron emission tomography studies. However, there are no single photon emission computed tomography or positron emission tomography data examining gender differences in CBF in healthy children. This study used TCD ultrasonography to provide some estimate of gender-related differences in CBF of the MCA and BA in young boys and girls. Data from the present study indicate that the observed adult gender differences in Vmca are present in young children before puberty. In addition, gender differences in Vbas are also present during early childhood, suggesting that whatever factor may be responsible for the higher CBFV in young girls affects both the anterior and posterior cerebral circulations. This information in healthy children is important to our understanding of changes in disease states, such as traumatic brain injury and diabetic kid-
There are several potential explanations for the higher Vmca and Vbas observed in females, irrespective of age. These include gender differences in blood viscosity because of hematocrit, hormones, vessel (MCA or BA) size, cerebral metabolism, and/or CVR. Identifying which of these factors is the major reason for the observed finding is challenging, however, because comparative data of these factors between boys and girls are largely lacking. First, whereas girls age 12 to 18 years have a lower hematocrit than boys,17,18 the 1 study to examine hematocrit in prepubertal boys and girls 4 to 8 years of age reported no gender differences.19 We did not have permission to obtain hematocrit in our healthy prepubertal subjects, but this should not be problematic, because there are no gender differences in hematocrit in children in this age group. Second, we discussed previously the role of hormones as an explanation for our findings of higher Vmca and Vbas in 10- to 16-year-old girls because of the following reasons: (1) estrogen and testosterone levels change during this time and vary by gender, (2) animal data have demonstrated that estrogens improve vasodilatation via an enhancement of endothelial nitric oxide synthase,20 and (3) testosterone and/or its metabolites increased cerebral perfusion in healthy adults.21 However, sex steroid levels are normally very low in boys and girls before puberty, and with a reference range for estradiol of <10 ng/mL, reference range for progesterone of <10 ng/dL, and reference range for progesterone of <0.7 ng/mL for both genders,22 it is unlikely that hormonal differences account for the observed gender differences in Vmca and Vbas during early childhood. Third, although it is thought that internal carotid artery and MCA diameters typically reach adult size by 6 years,23 details of the increase in vessel diameter by age and gender are not available. In 1 study of 156 cerebral angiograms of 133 adults and children (72 males and 84 females), Gabrielsen and Greiz24 reported no relationship between age and vessel size but described smaller internal carotid artery and MCA diameters in females compared with males. However, none of the angiograms were from children <10 years of age, and data from children were not analyzed separately. Therefore, it is unclear how cerebral vessel size influences the relationship between gender and CBF or CBFV in young children. Fourth, gender-related difference in Vmca and Vbas might be because of differences in cerebral metabolic rate (CMR).25 Although Kennedy and Sokoloff26 have shown that children have higher CMRs than adults, probably corresponding with the sharp rise in neuronal activity during this time,27 gender differences in CMR have not been examined. Finally, CVR may also impact CBF: Kennedy and Sokoloff26 also reported lower CVR in young children compared with adults. Describing similar findings, in 1988, Bode1 reported higher resistance indices (RIs) in newborns during the first days of life compared with after 1 year of life. After birth, RI slowly decreased reaching a nadir at 1 to 3 years of age before slightly increasing to adult levels between 10 and 16 years of age. However, gender was not considered, and he reported no difference in RI between the anterior and posterior cerebral circulations. In the present study, we estimated CVR from the MAP and CBFV and found that 4- to 8-year-old girls had lower CVR of BA than age-matched boys (Table 2). Review of our previous CVR data from older children shows that boys have higher CVR of both anterior and posterior cerebral

TABLE 3  CVRe of MCA and BA by Age and Gender in the Supine Position

<table>
<thead>
<tr>
<th>Age, y</th>
<th>CVRe of MCA, mm Hg · s/cm</th>
<th>P</th>
<th>CVRe of BA, mm Hg · s/cm</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys</td>
<td>Girls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4–8</td>
<td>0.78 ± 0.13 (0.60–1.07)</td>
<td>0.76 ± 0.13 (0.55–1.13)</td>
<td>.64</td>
<td>1.17 ± 0.19 (0.87–1.64)</td>
</tr>
<tr>
<td>10–16</td>
<td>1.11 ± 0.24 (0.69–1.56)</td>
<td>0.90 ± 0.28 (0.71–1.53)</td>
<td>.03</td>
<td>1.60 ± 0.36 (0.96–2.26)</td>
</tr>
</tbody>
</table>

Data in parentheses are 95% confidence intervals.

TABLE 4  Cerebral Autoregulation Data in 4- to 8-Year-Old Boys and Girls

<table>
<thead>
<tr>
<th>ARI</th>
<th>Boys</th>
<th>Girls</th>
<th>Mean ± SD</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCA</td>
<td>0.97 ± 0.06</td>
<td>0.94 ± 0.11</td>
<td>0.96 ± 0.09</td>
<td>.22</td>
</tr>
<tr>
<td>BAS</td>
<td>0.94 ± 0.13</td>
<td>0.94 ± 0.11</td>
<td>0.94 ± 0.12</td>
<td>.97</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>0.95 ± 0.11</td>
<td>0.94 ± 0.11</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>P</td>
<td>.01</td>
<td>.03</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

TABLE 5  Vmca and Vbas by Age and Gender in the Supine Position

<table>
<thead>
<tr>
<th>Age, y</th>
<th>Vmca, cm/s</th>
<th>Vbas, cm/s</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys</td>
<td>Girls</td>
<td></td>
</tr>
<tr>
<td>4–8</td>
<td>92 ± 13</td>
<td>99 ± 11</td>
<td>.03</td>
</tr>
<tr>
<td>10–16</td>
<td>75 ± 16</td>
<td>89 ± 16</td>
<td>.005</td>
</tr>
<tr>
<td>P</td>
<td>.002</td>
<td>.03</td>
<td>—</td>
</tr>
</tbody>
</table>
circulations than girls and that CVRe increases with age (Tables 3 and 6). Although we cannot separate the influence of cerebrovascular tone from vessel size on CVRe, these data suggest that a decrease in CVRe parallels the age- and gender-related increase in CBFV. We speculate that a higher CMR in girls might explain the noted lower CVRe and higher Vmca/Vbas in girls than in boys.

In a previous study, we reported greater ARIBas in girls compared with boys, whereas boys were found to have greater autoregulation of MCA. In this study of younger children, we did not find any difference in ARIBas of either the anterior or posterior cerebral circulations. There are 2 possible explanations for this difference. First, gender differences in cerebral autoregulation may be hormonally mediated and coincide with puberty. Recent reports show that mechanisms involved in cerebral autoregulation may vary as a function of age. For example, experimental evidence derived from newborn piglets shows the prominent role of prostaglandins in the control cerebral autoregulation in the newborn period. On the other hand, the nitric oxide system is thought to have a less prominent role in the newborn but increases in importance with increasing age. Alternatively, given the small size of both studies, both type I and type II errors may have occurred, and larger number of subjects may be needed to detect potential gender differences in cerebral autoregulation.

Cerebral autoregulation can be evaluated by measuring changes in CBFV in response to changes in blood pressure using either the static or dynamic method of testing. We used the static cerebral autoregulation test (tilting test), because it is noninvasive, does not involve pharmacological intervention, and is well tolerated by children. We attempted to examine Vmca, Vbas, and cerebral autoregulation in children 2.5 to 3 years of age, but the lack of subject cooperation precluded obtaining quality data in these children. Therefore, obtaining such data in very young children may only be feasible while they are either sedated or anesthetized.

There are some limitations to this study. First, we measured CBFV using TCD ultrasonography and not CBF. However, TCD ultrasonography has been used to estimate CBF in children with various medical diseases, such as sickle cell disease, and the use of TCD data has allowed us to better understand the effects of pathologic conditions on cerebrovascular physiology, such as the hyperemic state in pediatric traumatic brain injury. We could not control the blood pressure to ensure an exact drop in MAPE to precipitate an autoregulatory response. However, all of the study participants achieved or more than a 5-mm Hg drop in MAPE with position change from supine to upright. We did not evaluate cerebral arteries other than the MCA and BA. Although we did not monitor end-tidal CO2 and cannot exclude the effect of changes in ventilation on CBF, we found no difference in respiratory rate in each position between boys and girls. We could not control CMR or flow metabolism coupling in our awake study participants. However, we recorded measurement under a steady-state condition by minimizing distraction during the study. We calculated CVRe from Vmca and MAPE data. Finally, because there are no significant differences in hormone concentrations observed normally between prepubertal boys and girls, we did not obtain hormone levels from our study participants.

Our small study shows that prepubertal girls 4 to 8 years of age have higher Vmca and Vbas than age-matched boys. This difference may reflect either the influence of CMR and or CVRe on CBFV between the sexes. The lack of gender differences in ARIBas of both circulations during early childhood, in the presence of such differences in pubertal children, postulates a potential hormonal role in the control of cerebral autoregulation before adulthood. Our findings provide new information regarding gender differences in cerebrovascular physiology in healthy awake children. We have also demonstrated the feasibility of conducting such studies in children as young as 4 years of age.

### REFERENCES


### TABLE 6

<table>
<thead>
<tr>
<th>Position</th>
<th>Vmca, cm/s</th>
<th>Vbas, cm/s</th>
<th>CVRe of MCA, mm Hg·s·cm⁻¹</th>
<th>CVRe of BA, mm Hg·s·cm⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>4–8</td>
<td>95±12</td>
<td>66±11</td>
<td>0.77±0.13</td>
<td>1.11±0.19</td>
</tr>
<tr>
<td>10–16</td>
<td>82±17</td>
<td>50±12</td>
<td>1.0±0.26</td>
<td>1.45±0.38</td>
</tr>
<tr>
<td><em>p</em></td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

TONTISIRIN et al


Neurodevelopmental Consequences of Early Traumatic Brain Injury in 3-Year-Old Children

Heather T. Keenan, MDCM, PhD; Stephen R. Hooper, PhD; Crista E. Wetherington, PhD; Maryalice Nocera, MSN; Desmond K. Runyan, MD, DrPH

*Department of Pediatrics, University of Utah, Salt Lake City, Utah; †Clinical Center for the Study of Development and Learning, Department of Psychiatry; ‡School of Education, §University of North Carolina Injury Prevention Research Center, and ¶Departments of Social Medicine and Pediatrics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

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

ABSTRACT

OBJECTIVES. The purpose of this work was to determine cognitive and adaptive behavioral outcomes of children with traumatic brain injury acquired before age 2 years and to compare outcomes between inflicted versus noninflicted brain injury.

PATIENTS AND METHODS. All North Carolina children hospitalized in an ICU for a traumatic brain injury before age 2 years between the years 2000 and 2001 were eligible for study entry. A total of 112 surviving children were prospectively identified, 52 (46%) of whom had complete follow-up. Thirty-one control children were recruited from preschool settings. Control subjects were chosen to be demographically similar to case subjects. Child measures of cognition and adaptive behavior at age 3 years were measured and compared between children with and without traumatic brain injury and children with inflicted and noninflicted traumatic brain injury.

RESULTS. Sixty percent of injured children were >1 SD below normal on cognitive testing. Forty percent of injured children scored >1 SD below normal on adaptive behavior testing. Children with inflicted traumatic brain injury performed more poorly on tests of cognition and adaptive behavior. Glasgow Coma Scale ≥13, absence of seizures, income above twice the poverty guidelines, and high social capital were associated with improved outcomes. Injured children had lower scores than uninjured control children after adjustment for socioeconomic status.

CONCLUSIONS. Very young children with mild-to-severe traumatic brain injury as measured by the Glasgow Coma Scale are at risk for global cognitive deficits more than a year after the time of injury. Inflicted brain injury is associated with more severe injury and worse outcomes. This is less optimistic than findings in this same cohort 1 year after injury. Family characteristics seem to play a role in recovery after injury.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2313
doi:10.1542/peds.2006-2313

Dr Keenan had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Key Words
abuse, traumatic brain injury, shaken baby syndrome, injury, developmental disabilities

Abbreviations
TBI—traumatic brain injury
SIB-R—Scales of Independent Behavior-Revised
GCS—Glasgow Coma Scale
RR—relative risk
CI—confidence interval
IQR—interquartile range
OR—odds ratio

Accepted for publication Sep 27, 2006
Address correspondence to Heather T. Keenan, MDCM, PhD, 295 Chipeta Way, PO Box 581289, Salt Lake City, UT 84158. E-mail: heather.keenan@hsc.utah.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
TRAUMATIC BRAIN INJURY (TBI) is one of the most common causes of childhood disability in the United States, with a high proportion of injuries occurring in children <4 years old. Most longitudinal pediatric studies of the cognitive consequences of TBI have been performed in school-aged children; however, recent reports indicate that very young children may be more vulnerable to the deleterious effects of acquired TBI than older children. Longitudinal follow-up of very young children is important, because they must both regain skills and acquire new, more complex skills.

We previously recruited a cohort of children who experienced a TBI before 2 years of age. Approximately half of this cohort were victims of abuse. This study assessed the neurodevelopmental status of these children at 3 years of age. We examined the influence of injury mechanism, injury severity, and social domains on specific developmental outcome. We hypothesized that children who sustained an inflicted TBI before age 2 years would demonstrate more pervasive deficits in cognitive development and adaptive behavior compared with children with noninflicted TBI. We also expected that differences in the TBI group and the uninjured control subjects would persist after adjustment for socioeconomic disadvantage.

METHODS

TBI Cohort

All children <2 years old who were North Carolina residents and presented to a PICU after a TBI during the years 2000 and 2001 and survived were eligible for inclusion. Methods of recruitment and ascertainment of mechanism of injury have been described previously. Briefly, all of the children had either radiographic or pathologic evidence of a nonpenetrating intracranial injury. Mechanism of injury was decided by the child protection team at the hospital of origin and reviewed by 2 of the authors. The legal guardian was asked to consent to interview at 1 and 2 years postinjury. Families participating in telephone interviews were invited to have their child receive a neurodevelopmental evaluation. Four consenting families moved from the state before evaluation. The children of these 4 families did not have a Mullen score or anthropometric data collected, but all of the other family information and the Scales of Independent Behavior-Revised (SIB-R) were recorded.

Comparison Group

A comparison group of typically developing children, who had similar socioeconomic and racial backgrounds to the TBI group, was recruited from preschools in North Carolina and Southern Virginia. No comparison child had known head injury or other cognitive or physical disabilities by maternal report. This study was approved by the institutional review board at the University of North Carolina at Chapel Hill.

Procedures

Children were tested by 2 members of a team composed of 2 school psychology doctoral students and a pediatric nurse practitioner. The doctoral students were trained in the assessment instruments by a licensed neuropsychologist (Dr Hooper), who accompanied the team on the first 3 home visits and random subsequent visits to ensure consistency of the examinations. Family and child-specific data were collected from the maternal caregiver.

Family Level Data

Family data included whether the child was in his/her home of origin, total income to the household, number of people supported by that income, mother’s social capital, and the Hollingshead Four-Factor Index of Social Status (A. B. Hollingshead, PhD, Four Factor Index of Social Status, unpublished manuscript, 1975). Income level was compared with the published North Carolina poverty guidelines from the year 2000. Social capital is a construct incorporating a person’s integration with his/her community and family. High social capital has been associated with positive developmental outcomes in a group of children aged 2 to 5 years who were thought to be at high risk for poor outcomes secondary to an adverse social or economic environment. The social capital index was used as a bivariate descriptor with ≥4 considered “high” social capital as per the study by Runyan et al.

Child Level Data

Child injury data, including presenting modified pediatric Glasgow Coma Scale (GCS), presence of posttraumatic seizures, and receipt of cardiopulmonary resuscitation, was ascertained at the time of injury by chart review. Children were grouped by GCS score, with mild injury defined as a GCS of 13 to 15, moderate injury as a GCS of 9 to 12, and severe as a GCS of 3 to 8. Children were examined for major disabilities, use of adaptive aids, and anthropometric data at the time of the 3-year-old evaluation.

Cognitive-Developmental Evaluation

The Mullen Scales of Early Learning was chosen to assess cognitive abilities and developmental status. Four Mullen subscales were used: visual reception, fine motor, expressive language, and receptive language. An overall composite index (mean: 100; SD: ±15), the early learning composite (composite), was calculated with its standard score. The Mullen has a sufficiently low floor to assess the most impaired child in the sample. For our purposes, the Mullen composite was dichotomized into reference range and greater (≥85) and below reference range.
Adaptive Behavior Evaluation

The SIB-R reflected the maternal caregiver’s perception of the child’s adaptive behavior. This age-normed scale assesses skills needed to function independently in age-appropriate settings. As above, the scale was dichotomized into reference range (≥85) or below reference range.

Statistical Analysis

Characteristics of eligible families of injured children who did not participate were compared with participants using χ² statistics for categorical variables. Medians were calculated for nonnormally distributed data and compared with the Wilcoxon rank sum test. P values of .05 were considered statistically significant throughout.

Comparison of Children With Inflicted and Noninflicted TBI

Injury and family characteristics were examined. The relative risk (RR) of a low GCS (<13 vs ≥13), dichotomized child race, social capital index, and income status dependent on injury mechanism was calculated with 95% confidence intervals (CIs). The Mullen composite, each Mullen subscale, and the SIB-R were compared. The distributions of the 2 groups were examined by categorizing them into the number of SDs below normal. Then, as the scales were nonnormally distributed, the median and interquartile range (IQR) were calculated for each scale, and the Wilcoxon rank sum test was used to compare scales by injury mechanism. Finally, a correlation coefficient was calculated using the Mullen composite and SIB-R to test whether or not cognitive outcome and adaptive behavior were related.

Predictors of poor outcome were assessed with a bivariate analysis using child and family covariates as predictors and the dichotomized Mullen composite as the outcome. The RR of poor outcome with 95% CI was calculated for categorical variables. Statistically significant covariates in the bivariate models were entered into a proportional odds model. The proportional odds model is a multivariate model used to examine the odds of a child performing 1 SD better on the Mullen composite per unit of change in the predictor variables.

Comparison of Injured Children to Uninjured Control Subjects

Families of injured and uninjured children were compared using the χ² test. Means were calculated for the Mullen composite, Mullen subscales, and the SIB-R. Student’s t test was used to compare the mean scores of injured to uninjured children. The odds of injured versus uninjured children falling below norm on the Mullen composite were calculated.

To insure that differences observed between the TBI and uninjured groups were not because of socioeconomic status, a propensity score was created using the Hollingshead Index, family income, and child gender. A common odds ratio (OR) adjusted for propensity score was calculated using exact methods.

RESULTS

Of the 112 surviving children identified with a TBI, 72 (64.3%) participated in telephone follow-up. Fifty-two (72.2%) of the 72 children followed by telephone received a 3-year-old evaluation. Therefore, 46.4% of all of the eligible children in North Carolina were evaluated. Eligible nonparticipants were similar to children and families who participated in the home visit at initial enrollment (Table 1).

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Baseline Characteristics of Eligible Families and Children Who Did Not Complete the Home Visit Compared With Those Who Completed the Home Visit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>Did Not Complete Home Visit (N = 60)</td>
</tr>
<tr>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Children</td>
<td></td>
</tr>
<tr>
<td>Inflicted</td>
<td>35</td>
</tr>
<tr>
<td>Male</td>
<td>37</td>
</tr>
<tr>
<td>White</td>
<td>30</td>
</tr>
<tr>
<td>GSC</td>
<td></td>
</tr>
<tr>
<td>13–15</td>
<td>29</td>
</tr>
<tr>
<td>9–12</td>
<td>13</td>
</tr>
<tr>
<td>3–8</td>
<td>16</td>
</tr>
<tr>
<td>Missing</td>
<td>2</td>
</tr>
<tr>
<td>Age at injury, median (IQR), mo</td>
<td>5.0 (2.0–10.0)</td>
</tr>
<tr>
<td>Mothers*</td>
<td></td>
</tr>
<tr>
<td>Age at child’s injury, median, IQR, y</td>
<td>24 (20–28)</td>
</tr>
<tr>
<td>Married</td>
<td>25</td>
</tr>
<tr>
<td>Education: high school or more</td>
<td>36</td>
</tr>
</tbody>
</table>

No statistically significant differences exist between the 2 groups.
* These data are from the child’s home of origin.
Injured Group: Family and Child Characteristics

Injured children were most frequently cared for by families with 2 parents (63.5%) whose maternal and paternal caregivers had at least a high school education (78.8% and 77.8%, respectively) and were employed (84.6%). Approximately 29% of the maternal caregivers were foster or adoptive parents. Families were generally poor; 30% were below the North Carolina poverty guidelines, and 69% were below twice the North Carolina poverty guidelines. The most frequent source of maternal income was work performed by herself or another adult (86.5%) as opposed to child support or government support. The mean Hollingshead Index was 32.4 (SD: 13.6). Most families had a social capital index of ≥4 (63.5%).

The median child age at injury was 4.2 months (IQR: 1.8–9.9). Twenty-seven children (52%) had inflicted TBI. One child with inflicted TBI was premature (32 weeks’ gestational age). No child had a history of congenital cardiac, neurologic, or pulmonary disease. No child sustained a second TBI during follow-up.

When children were compared by mechanism of injury, more children with inflicted TBI had a GCS <13 than children with noninflicted TBI (55% vs 24%, respectively; RR: 1.9; 95% CI: 1.1–3.1). Children were similar when compared by race (RR: 0.8; 95% CI: 0.5–1.4), age at injury (3.7 vs 6.8 months; P = .2), and age at evaluation (3.1 vs 3.2 years; P = .5). Families caring for children after inflicted TBI were not substantively different from families caring for children with noninflicted injuries by Hollingshead Index (P = .8), social capital (RR: 1.7; 95% CI: 0.9–3.2), or income below the North Carolina poverty guidelines (RR: 0.6; 95% CI: 0.3–1.3) or twice below the poverty guidelines (RR: 1.0; 95% CI: 0.6–1.7).

Anthropometrics and Disabilities

The children’s physical assessment revealed that 8.2% were <5th percentile for height and weight. Nearly a quarter of children (23%) were ≥2 SDs below normal for head circumference. The most frequent disabilities were delay in speech acquisition in 19 (36.5%), three quarters of whom had inflicted injuries. Eleven children (21.1%) had ongoing seizure disorders (73% inflicted TBI). Other disabilities included blindness (3), spasticity (5), quadriplegia (1), and hemiparesis (11). Mobility problems were frequent; 17 children (32.7%) required wheel chairs, walkers, or braces (76.4% inflicted TBI). Other adaptive aids included glasses (4), hearing aids (1), gastric feeding tubes (4), and tracheostomy (1).

Neurodevelopmental Evaluation

Children with TBI scored below population norms on the Mullen composite and all of its subscales but scored within population norms in adaptive behavior (Table 2). The composite score was examined by the child presenting modified GCS. Children with a GCS <13 were at an increased risk (RR: 6.6; 95% CI: 1.7–25.5) of having a composite below reference range, although a GCS ≥13 did not guarantee a normal score. In fact, 37% of children with a GCS ≥13 scored below normal on the composite (Fig 1). When the Mullen composite was examined by injury mechanism, more children with inflicted TBI fell into the lowest category (>3 SD below normal) compared with those with noninflicted TBI (40% vs 4.3%; RR: 2.6; 95% CI: 1.6–4.2; Fig 2). Children with inflicted injuries did more poorly across all of the Mullen subscales (Table 2).

Children with inflicted and noninflicted TBI were compared on the SIB-R. No statistical difference in means was seen between groups (P = .2); however, the distributions were different (Table 2). Children with inflicted TBI were more likely to be ≥3 SDs below normal than children with noninflicted TBI (RR: 1.6; 95% CI: 1.0–2.6) on the SIB-R (Fig 3). The SIB-R was moderately correlated with the Mullen composite (R² = 0.6).

Child and social covariates associated with outcome on the Mullen composite included male gender, posttraumatic seizures, GCS <13, social capital index, and poverty status (Table 3). The multivariate model confirmed these results. After adjustment for all covariates in the model, high GCS category (OR: 12.1; 95% CI: 3.0–48.9), absence of seizures (OR: 6.1; 95% CI: 1.6–24.1), family income above twice the poverty guidelines

TABLE 2
Cognitive and Adaptive Behavior Outcomes of Children With Inflicted and Noninflicted TBI

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Injured Children (N = 48)</th>
<th>Inflicted (n = 25)</th>
<th>Noninflicted (n = 23)</th>
<th>Pα</th>
<th>Norms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median (IQR)</td>
<td>Median (IQR)</td>
<td>Median (IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual reception subscale</td>
<td>35.0 (20.0–46.5)</td>
<td>29.0 (20–43)</td>
<td>38.0 (30–54)</td>
<td>.04</td>
<td>50 ± 10</td>
</tr>
<tr>
<td>Fine motor subscale</td>
<td>36.0 (20–52)</td>
<td>20.0 (20–45)</td>
<td>45.0 (28–53)</td>
<td>.02</td>
<td>50 ± 10</td>
</tr>
<tr>
<td>Receptive language subscale</td>
<td>38.5 (24–48)</td>
<td>24.0 (20–46)</td>
<td>43.5 (35–51)</td>
<td>.01</td>
<td>50 ± 10</td>
</tr>
<tr>
<td>Expressive language subscale</td>
<td>34.5 (24.5–43.0)</td>
<td>30.0 (20–43)</td>
<td>38.3 (31–45)</td>
<td>.02</td>
<td>50 ± 10</td>
</tr>
<tr>
<td>Early learning comprehension</td>
<td>77.0 (57.0–91.5)</td>
<td>68.0 (49–86)</td>
<td>84.0 (68–100)</td>
<td>.02</td>
<td>100 ± 15</td>
</tr>
<tr>
<td>SIB-R standard score</td>
<td>97.0 (65–120)b</td>
<td>94.0 (12–122)c</td>
<td>100.0 (79–113)d</td>
<td>.2</td>
<td>100 ± 15</td>
</tr>
</tbody>
</table>

α Wilcoxon Rank Sum Test comparing inflicted to noninflicted traumatic brain injury.

b N = 52.

c n = 27.
d n = 25.

PEDIATRICS Volume 119, Number 3, March 2007 e619
(OR: 14.5; 95% CI: 2.2–93.5), and high social capital (OR: 5.0; 95% CI: 1.3–18.9) were all predictors of 1 level better of outcome; however, estimates were imprecise because of small numbers.

Comparison of Children With TBI to Uninjured Children

Thirty-one uninjured families and children participated. The mean age of child evaluation was 3.6 years (SD: 0.3 years). Families of the uninjured and TBI children had similar characteristics when compared by marital status ($P = .2$), employment status ($P = .5$), maternal and paternal education of high school or more ($P = .2$ and $.7$, respectively), income source ($P = .5$), and percentage below the poverty guidelines ($P = .8$). Families of uninjured children had a significantly higher Hollingshead Index (median: 42.1; SD: 14.4; $P = .002$) compared with families of injured children, and uninjured children were more likely to live with their biological mother ($P = .01$).

Scores of the injured and uninjured children overlapped on both the Mullen scales and the SIB-R. However, TBI children scored significantly lower on the composite and 3 of the 4 Mullen subscales than uninjured children. Both groups scored poorly on the fine-motor subscale (Table 4). The odds of scoring below the reference range on the Mullen composite for a child with TBI compared with an uninjured child was 4.9 (95% CI: 1.9–13.3). After adjustment for propensity score, the common OR of scoring below the reference range remained at 3.9 (95% CI: 1.3–12.4) for children with TBI. Children with TBI also scored lower on the SIB-R ($P = .001$) than uninjured children.

DISCUSSION

This study found that children who suffer a TBI before age 2 years are at high risk for pervasive cognitive deficits and deficits in adaptive behavior. The cognitive deficits are global and include problems in motor, visual processing, and receptive and expressive language. These deficits persist when controlled for social status. Children with inflicted TBI had greater disability and more severe cognitive and adaptive behavior problems than children with noninflicted TBI. All of the participants with TBI were tested ≥1 year postinjury; thus, all should have completed their most rapid recovery phase. The results of this formal neurocognitive testing performed, on average, 2 years after injury provide a less optimistic picture than results from telephone follow-up performed for this same cohort of children at 1 year postinjury. At that time, children with inflicted injury fared more poorly than children with noninflicted injury; however, more than half of the cohort had, at most, mild deficits recognized. The more optimistic picture obtained from parental interview 1 year after injury...
could reflect the limitations of the tools used to assess the children’s cognitive function or may reflect parents’ optimism to having their child make some cognitive gains after a critical illness. To our knowledge, these data represent the longest prospective neurodevelopmental follow-up of children with inflicted and noninflicted TBI acquired at similar ages.

Children at risk for cognitive deficits included children with “mild” injury severity as measured by the GCS. Thirty-seven percent of children with a mild GCS score12–14 and all of the children with a moderate range GCS tested below population norms. Overall, the children’s cognitive abilities were paralleled by their parent-rated adaptive behavior scores; however, cognitive deficits were identified in a subset of children rated well on scales of adaptive behavior.

### Table 3: Bivariate Analysis Showing the RR of Being >1 SD Below Normal on Mullen Early Learning Composite Score for Injured Children

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>&gt;1 SD Below</th>
<th>Reference Range</th>
<th>RR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Physiologic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>20</td>
<td>69.0</td>
<td>7</td>
<td>36.8</td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
<td>31.0</td>
<td>12</td>
<td>63.2</td>
</tr>
<tr>
<td>Mechanism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflicted</td>
<td>17</td>
<td>58.6</td>
<td>8</td>
<td>42.1</td>
</tr>
<tr>
<td>Noninflicted</td>
<td>12</td>
<td>41.4</td>
<td>11</td>
<td>57.9</td>
</tr>
<tr>
<td>Seizure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>17</td>
<td>58.6</td>
<td>5</td>
<td>16.3</td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>41.4</td>
<td>14</td>
<td>73.7</td>
</tr>
<tr>
<td>Cardiopulmonary resuscitation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>7</td>
<td>24.2</td>
<td>2</td>
<td>10.5</td>
</tr>
<tr>
<td>No</td>
<td>22</td>
<td>75.9</td>
<td>17</td>
<td>89.5</td>
</tr>
<tr>
<td>GCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–12</td>
<td>19</td>
<td>65.5</td>
<td>2</td>
<td>10.5</td>
</tr>
<tr>
<td>13–15</td>
<td>10</td>
<td>34.5</td>
<td>17</td>
<td>89.5</td>
</tr>
<tr>
<td>Social</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social capital index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;4</td>
<td>14</td>
<td>48.3</td>
<td>3</td>
<td>15.8</td>
</tr>
<tr>
<td>≥4</td>
<td>15</td>
<td>51.7</td>
<td>16</td>
<td>84.2</td>
</tr>
<tr>
<td>Below the poverty level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>12</td>
<td>42.9</td>
<td>2</td>
<td>10.5</td>
</tr>
<tr>
<td>No</td>
<td>16</td>
<td>57.1</td>
<td>17</td>
<td>89.5</td>
</tr>
<tr>
<td>Twice below the poverty level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>25</td>
<td>89.3</td>
<td>9</td>
<td>47.4</td>
</tr>
<tr>
<td>No</td>
<td>3</td>
<td>10.7</td>
<td>10</td>
<td>52.6</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>7</td>
<td>24.1</td>
<td>4</td>
<td>21.1</td>
</tr>
<tr>
<td>High school or greater</td>
<td>22</td>
<td>75.9</td>
<td>15</td>
<td>78.9</td>
</tr>
<tr>
<td>Hollingshead Index (categorized)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More than mean</td>
<td>18</td>
<td>69.2</td>
<td>8</td>
<td>44.4</td>
</tr>
<tr>
<td>Less than mean</td>
<td>8</td>
<td>30.8</td>
<td>10</td>
<td>55.6</td>
</tr>
</tbody>
</table>

### Table 4: Cognitive and Adaptive Behavior Outcomes of Uninjured Children and Children With TBI

<table>
<thead>
<tr>
<th>Variable</th>
<th>Uninjured children (N = 31), Mean ± SD</th>
<th>All Injured Children (N = 48), Mean ± SD</th>
<th>P*</th>
<th>Norms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mullen scale scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual reception subscale</td>
<td>50.9 ± 11.9</td>
<td>36.4 ± 14.6</td>
<td>.0001</td>
<td>50 ± 10</td>
</tr>
<tr>
<td>Fine motor subscale</td>
<td>41.0 ± 13.2</td>
<td>37.8 ± 17.1</td>
<td>.25</td>
<td>50 ± 10</td>
</tr>
<tr>
<td>Receptive language subscale</td>
<td>49.0 ± 12.4</td>
<td>37.8 ± 14.1</td>
<td>.0005</td>
<td>50 ± 10</td>
</tr>
<tr>
<td>Expressive language subscale</td>
<td>45.4 ± 8.9</td>
<td>35.3 ± 11.8</td>
<td>.0001</td>
<td>50 ± 10</td>
</tr>
<tr>
<td>Early learning comprehension</td>
<td>940 ± 18.7</td>
<td>77.3 ± 22.0</td>
<td>.0007</td>
<td>100 ± 15</td>
</tr>
<tr>
<td>Scales of Independent Behavior Standard Score-Revised</td>
<td>116 (107–131)‡</td>
<td>97 (65–120)‡</td>
<td>.001‡</td>
<td>100 ± 15</td>
</tr>
</tbody>
</table>

* Data are from a t test comparing uninjured with injured children.

b Data are median (IQR).

Data are from a Wilcoxon rank sum test comparing uninjured with injured children.

at Prince Of Songkla Univ on April 19, 2007 www.pediatrics.orgDownloaded from
Children with inflicted TBI had worse outcomes than children with noninflicted TBI; however, degree of disability was associated more strongly with injury severity than injury mechanism. The disproportionate burden of injury in children with inflicted TBI may result from a delay in receiving care, because many children with inflicted injuries become recognized only when they have respiratory distress, seizures, or are unarousable.\(^{14}\) This delay in recognition could add a secondary brain insult, causing the child to have a worse outcome than children with immediately recognized injury. Another possibility is that mechanisms of injury causing inflicted TBI result in more severe injury than those typical of noninflicted TBI. The difference in outcome does not seem to be from social factors alone, because the children’s homes of origin were similar in the 2 groups.\(^{14}\)

These results agree partially with a previous study of young children with noninflicted TBI which reported that preschool children with severe TBI had cognitive deficits across multiple domains at follow-up.\(^{19}\) However, unlike our study, children with mild-to-moderate injury did not show persistent deficits. The difference in results may be secondary to older age at injury, higher socioeconomic status, or differences in measures of injury severity in the comparison study.

The severe delay in children with inflicted injuries is consistent with previous reports.\(^{20,21}\) A previous prospective study comparing children with inflicted to noninflicted amnesia found that 5% of children with noninflicted TBI compared with 45% of children with inflicted TBI scored in the mentally deficient range on formal cognitive testing.\(^{21}\) A combined prospective (\(n = 12\)) and cross-sectional (\(n = 13\)) study of 25 children with inflicted TBI found that 68% of survivors were abnormal on follow-up.\(^{20}\) These previous studies have been hampered by a lack of prospective longitudinal follow-up,\(^{20,22}\) differential age at injury dependent on injury type,\(^{21}\) and length of follow-up. No previous study has used socioeconomically similar controls.

Environmental influences including poverty and social capital played a role in recovery from injury. Families with better financial means and those able to access family or community resources tended to do better. This finding is unlikely to be secondary to access to medical care, because most families could identify a primary pediatrician and access ancillary therapies.\(^{4}\) Social and economic disadvantage have been shown to be important in infant cognitive development.\(^{23-25}\) Preterm infants have shown short-term developmental gains with the provision of increased stimulation and/or increased social support to families.\(^{26}\) As strategies to enhance parent-infant interactions are teachable to parents, this may be a malleable recovery factor.

This study has limitations. The primary limitation is the size of the cohort. Although the study size limits some conclusions that can be drawn from the analysis because of imprecision, it is the largest study of its kind and represents an effort to follow every child with early TBI during a defined period in an entire state. We were unable to obtain participation from all of the injured children limiting the study’s generalizability. Although this may introduce bias, our sample was similar in those characteristics measured compared with all of the children eligible for study. Finally, although data from the uninjured controls were adjusted for socioeconomic disadvantage, they may differ by important unmeasured covariates.\(^{27}\)

**CONCLUSIONS**

Children who acquire radiographically evident TBI before age 2 years have persisting deficits in both cognitive development and adaptive behavior. Cognitive delays were found both in children with mild injury and caregiver-rated normal adaptive behavior. Because a complete neurodevelopmental evaluation is not a routine part of care after early TBI, delays in cognition may not be identified unless formal comprehensive testing is performed.

**ACKNOWLEDGMENT**

Dr Keenan is supported by a grant from the National Institute for Child Health and Human Development (K23 HD041040-01A2).

**REFERENCES**


Weight Status in Young Girls and the Onset of Puberty

Joyce M. Lee, MD, MPH\textsuperscript{a,b}, Danielle Appugliese, MPH\textsuperscript{c}, Niko Kaciroti, PhD\textsuperscript{d}, Robert F. Corwyn, PhD\textsuperscript{e}, Robert H. Bradley, PhD\textsuperscript{e}, Julie C. Lumeng, MD\textsuperscript{d,f}

\textsuperscript{a}Pediatric Endocrinology, \textsuperscript{b}Child Health Evaluation and Research Unit, and \textsuperscript{c}Child Behavioral Health, Department of Pediatrics, and \textsuperscript{d}Center for Human Growth and Development, University of Michigan, Ann Arbor, Michigan; \textsuperscript{e}Data Coordinating Center, Boston University School of Public Health, Boston, Massachusetts; \textsuperscript{f}Center 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. We sought to examine the association between weight status in early childhood and onset of puberty.

PATIENTS AND METHODS. The study included 354 girls from the National Institute of Child Health and Human Development Study of Early Child Care and Youth Development. Girls were followed longitudinally with height and weight measurements at 36 and 54 months and grades 1, 4, 5, and 6 and with assessment of pubertal stage by physical examination and maternal report in grades 4 through 6. The main outcome was the presence of early puberty, indexed as follows: (a) breast development at or more than Tanner stage 2 by physical examination at grade 4; (b) breast development at or more than Tanner stage 3 by physical examination at grade 5; (c) maternal report of breast development at or more than Tanner stage 3 at grade 5; and (d) maternal report of menarche having already occurred (yes versus no) at grade 6. Multiple logistic regression models predicting early versus late puberty were constructed by using the covariate BMI \textit{z} score at 36 months, rate of change of BMI between 36 months and grade 1, race, maternal education, and maternal age of menarche.

RESULTS. BMI \textit{z} score at 36 months, rate of change of BMI between 36 months and grade 1, an earlier age of maternal menarche, and nonwhite race were each consistently and positively associated with an earlier onset of puberty across the various measures of puberty.

CONCLUSIONS. Higher BMI \textit{z} score in girls as young as 36 months of age and higher rate of change of BMI between 36 months old and grade 1, a period well before the onset of puberty, are associated with earlier puberty, which suggests that increasing rates of obesity in the United States may result in an earlier average age of onset of puberty for US girls.
A number of studies have evaluated the association between weight status and onset of puberty, because excess adiposity or body fatness is hypothesized to be a causal factor for early pubertal development in girls. Cross-sectional studies have shown that pubertal girls have higher age- and gender-adjusted BMIs than their similarly aged prepubertal counterparts. However, it is unclear whether increased body fatness precedes puberty or vice versa, given that puberty is a period associated with marked changes in body composition, including increases in fat mass and weight gain.

The impact of weight status on puberty is a question of considerable importance, given that rates of obesity among children in the United States have doubled over the last 2 decades. Studies suggest that girls are entering puberty at younger ages compared with 30 years ago, and it has been hypothesized that the problem of childhood obesity is contributing to this secular trend.

Longitudinal studies are, therefore, needed to evaluate the directionality of the relationship between weight and onset of puberty. One recently published longitudinal study demonstrated that girls with a higher percentage of body fat at 5 years old and girls with larger increases in percentage of body fat from 5 to 9 years of age were more likely to have pubertal development by age 9 years. These results offered evidence that increased weight status precedes and may, therefore, be causally linked to an earlier onset of puberty. However, given that the period between ages 5 and 9 years encompassed the onset of puberty, the directionality of the relationship between puberty and body fat may have been confounded, because studies have shown that onset of puberty is associated with increases in BMI. We therefore sought to examine the association between weight status earlier than 5 years old and onset of puberty, using a longitudinal study design in a slightly larger sample of girls from diverse socioeconomic status drawn from 10 regions in the United States.

Methods

The National Institute of Child Health and Human Development Study of Early Child Care and Youth Development is a longitudinal study of child behavior and development, particularly related to child care experiences. Data from this study are available to qualified researchers by application. Parents and children were recruited through hospital visits to mothers shortly after the birth of a healthy singleton child in 1991 using a conditional random sampling plan. Details of the recruitment methods and sampling plan are available elsewhere.

Families lived in the areas of Little Rock, Arkansas; Irvine, California; Lawrence, Kansas; Boston, Massachusetts; Morganton, North Carolina; Philadelphia, Pennsylvania; Pittsburgh, Pennsylvania; Charlottesville, Virginia; Seattle, Washington; and Madison, Wisconsin. This study was approved by the institutional review boards of all of the pertinent institutions.

Anthropometric Data

The main predictor of interest was BMI before the onset of puberty. Height and weight were measured during laboratory visits at 36 months, 54 months, and grades 1, 4, 5, and 6 by trained research assistants. BMI and BMI z score for each age were calculated using the 2000 Centers for Disease Control and Prevention growth charts. Girls were classified as "at risk for overweight" if age- and gender-specific BMI was ≥85th percentile and <95th percentile and "overweight" if BMI was ≥95th percentile.

Puberty Measures

The outcome of interest was the presence of puberty, which we indexed as a dichotomous variable (“earlier puberty” versus “later puberty”). Measures of pubertal development were obtained at 3 annual laboratory visits occurring when girls were in grades 4 through 6, when girls were (mean ± SD) 9.6 ± 0.1, 10.6 ± 0.2, and 11.6 ± 0.1 year of age, respectively. Tanner staging at the laboratory visit physical examinations was performed by visual inspection by pediatric endocrinologists or nurse practitioners with experience in pubertal staging, using the methods of Herman-Giddens et al and others.

Mothers reported their daughter’s pubertal development at each of the 3 laboratory visits by selecting the drawing of the 5 stages of Tanner breast development that they felt was closest to their child’s development. In addition, at each of the visits, mothers responded to the question, “Has [your daughter] had her first menstrual period?” (yes versus no).

We indexed the presence of early puberty through several different measurements of breast development, as well as the onset of menses, to enhance the validity of our findings. Early puberty was defined in the following 4 ways: (a) breast development at or more than Tanner stage 2 by physical examination at the grade 4 visit; (b) breast development at or more than Tanner stage 3 by physical examination at the grade 5 visit; (c) maternal report of breast development at or more than Tanner stage 3 at the grade 5 visit; and (d) maternal report of menarche having already occurred (yes versus no) at the grade 6 visit.

The first definition of puberty used in this study was breast development at or more than Tanner stage 2 by physical examination at the grade 4 visit, representing the textbook definition of onset of puberty in girls. This measure has been used to define puberty in previous studies and provided a maximal sample size (n = 354). Because physical examination did not include breast palpation in this study, excess fat could be mistaken for breast development in overweight girls, leading to misclassification of prepubertal girls as being in puberty.
Therefore, an alternate definition of puberty, breast development at or more than Tanner stage 3 by physical examination at the grade 5 visit was also assessed, which may be less prone to misclassification, but also resulted in a smaller sample size \( (n = 292) \). Some, but not all, studies\(^{21} \) have documented good agreement between physician- and subject-reported assessments of puberty, with a \( \kappa \) coefficient of 0.81\(^{22} \) or a correlation of \( r = 0.87 \)\(^{23} \) for female breast development. Therefore, maternal report of breast development at or more than Tanner stage 3 at the grade 5 visit \( (n = 328) \) was used to define earlier puberty. Cohen’s \( \kappa \) for agreement of maternal report and physical examination for assessing at-or-more-than Tanner stage 3 development at grade 5 was 0.61 (95% confidence interval: 0.51–0.71). Finally, maternal report of menarche (yes versus no) at the grade 6 visit \( (n = 338) \) was used as a definition of earlier puberty. We chose to use the grade 6 maternal report of menses, because there was less missing information for the maternal report than for the grade 6 pubertal examination. In addition, assessment of Tanner staging in grade 6 would have represented a method that was similar to assessment of Tanner staging in grades 4 and 5, and we wished to use a different method for determining earlier and later onset of puberty.

**Covariates**

Mother-reported maternal age at menarche was included as a covariate given that later pubertal development in girls is associated with a maternal history of later pubertal development.\(^{24} \) Studies have also shown that higher socioeconomic status is associated with earlier pubertal development in girls;\(^{24} \) therefore, the income-to-needs (ITN) ratio at 36 months old was included as a continuous variable. The ITN ratio represents total family income relative to the poverty line for a family of a particular size; families with an ITN ratio <1 are considered “poor.” Maternal education, represented by the highest grade level of education completed by the mother, was a covariate, given that higher parental education in 1 previous study has been associated with earlier timing of puberty.\(^{25} \) Finally, race was included as a covariate (dichotomized into the categories “white” and “nonwhite”) given that black (compared with white) race is associated with earlier onset of puberty.\(^{26,27} \)

The initial sample in the study was composed of 1364 children (both boys and girls) and was representative of the demographics of the catchment areas from which the sample was recruited.\(^{11} \) By grade 6, 1077 families remained in the study.\(^{28} \) The sample for this analysis was restricted to girls with complete data for BMI at 36 months, 54 months, and grade 1; maternal age at menarche; and the pubertal outcome for each of the analyses. For the analysis for which the outcome was the presence of puberty based on physical examination Tanner staging at grade 4, complete data were available for 354 girls. The sample with complete data for BMI and puberty included in this analysis \( (n = 354) \) was compared with the sample with BMI but without puberty data \( (n = 92) \) by \( \chi^2 \) tests of association for categorical variables and \( t \) tests for continuous variables. Girls did not differ by race \( (P = .7) \), ITN ratio \( (P = .4) \), maternal education \( (P = .8) \), or BMI at grade 1 \( (P = .9) \).

Complete information on BMI, maternal age at menarche, and pubertal outcomes at later grades was available for a subset of the sample of 354 girls: 292 girls with a grade 5 physical examination, 328 girls with maternal report of Tanner stage 3 development in grade 5, and 338 girls with maternal report of menses in grade 6.

**Analysis**

All of the data analysis was performed by using SAS 9.1 (SAS Institute Inc, Cary, NC). Separate statistical analyses were conducted for each of the 4 different outcome measures of puberty. Bivariate statistics were performed to evaluate differences between girls who had earlier versus later puberty, and unadjusted analyses using logistic regression were performed to assess the association between BMI \( z \) score at each visit (36 months, 54 months, and grades, 1, 4, 5, and 6) and the odds of having earlier puberty.

The association between weight and early puberty could be related to the absolute BMI \( z \) score, the rate of change of BMI, or rapid increases in BMI early in childhood. Second-order polynomial trend scores were, therefore, created, using raw BMI scores rather than BMI \( z \) scores, because studies in similarly aged children suggest that raw BMI is superior to \( z \) scores for assessing changes in adiposity.\(^{28} \) The first-order polynomial term assesses the rate (slope) of change on BMI between 36 months and grade 1. The second-order polynomial term, named accelerated BMI, assesses the rate of change in BMI growth between 36 months and grade 1. This term was calculated by subtracting the actual BMI at 54 months from the expected linear growth of BMI at 54 months. Expected linear growth of BMI at 54 months was determined by predicting BMI at 54 months based on the calculated slope of BMI change between 36 months and grade 1. Values close to 0 would indicate linear growth, and large values would indicate a nonlinear/quadratic pattern of increases in BMI with growth that is accelerated over this period, indicating a faster rate of gain in BMI between 54 months and grade 1.

The unadjusted associations between rate of change of BMI and accelerated BMI and the odds of having earlier puberty were evaluated. Because polynomial trend scores are orthogonal, multiple logistic regression models for predicting earlier puberty were then performed, including the covariates of BMI \( z \) score at 36 months old; rate of change of raw BMI between 36 months and grade 1; and the acceleration in change in BMI, race, maternal education, and maternal age at
menarche. Because we did not find significant differences in the ITN ratio between girls with earlier and later puberty through our bivariate analyses, it was not included, to allow for the most parsimonious model. We evaluated for interactions between race and BMI z score, but these were not significant (data not shown).

**RESULTS**

Table 1 shows the demographic characteristics of the cohort overall and stratified by earlier and later puberty at the grade 4 visit. Most of the children were white, and of the nonwhite children (n = 63), 46 (73%) were black. Sixteen percent of the sample had an ITN ratio <1, indicating that they were living in poverty. In unadjusted bivariate analyses, earlier puberty was associated with nonwhite race, lower maternal education, earlier maternal age at menarche, and greater prevalence of child overweight and at risk for overweight. There were no significant differences in ITN ratio by puberty status.

Table 2 shows mean BMI, BMI z score, and weight status at each data collection point, as well as the Tanner stage in grades 4 through 6. By grade 4, 30% of the girls were either at risk for overweight or overweight. There were 168 girls who were considered “in puberty” (defined as Tanner stages 2, 3, 4, or 5) at grade 4. Of these girls, 139 were Tanner stage 2. Of the 139 girls, 46 progressed to stage 3 puberty by grade 5, and of those 46 girls, 23 were reported as having the first menstrual period by grade 6.

The results of unadjusted logistic regression models evaluating the association between BMI z score at 36 months through grade 6 and the odds of having earlier puberty using 4 different measures of puberty are shown in Table 3. Starting at 36 months of age, higher BMI z score at all of the ages was strongly associated with an earlier onset of puberty. For example, for each 1-point score at 36 months, the odds of having earlier puberty, as measured by Tanner stage 2 breast development at grade 4, increased by 44%. Similar findings were found across all of the measures of puberty. Rate of BMI change between 36 months and grade 1 and accelerated BMI were also strongly associated with an earlier onset of puberty.

The multiple logistic regression models predicting earlier puberty for each of the 4 puberty measures are presented in Table 4. Although the predictors did not reach statistical significance in every model, the patterns of association were similar such that overall, BMI z score at 36 months, rate of change of BMI between 36 months and grade 1, an earlier age of maternal menarche, and nonwhite race emerged as significant predictors of earlier puberty across the 4 models. Acceleration in BMI and maternal education were not significantly associated with earlier onset of puberty.

**DISCUSSION**

To our knowledge, this represents the first longitudinal study to report that higher BMI z score in girls as young as 36 months of age and a higher rate of change of BMI between 36 months old and grade 1, a period well before the onset of puberty, are associated with earlier puberty, even after adjustment for race, age of maternal menarche, and maternal education. The findings of this study provide additional evidence that increased body fatness precedes the onset of puberty in healthy girls.

A number of studies have suggested a causal link between excess body weight and an earlier timing of puberty. Frisch and McArthur1 first hypothesized in 1974 that a “critical body weight” was necessary for the onset of puberty. Support for this hypothesis includes animal studies that have demonstrated that restricting weight gain delays the timing of puberty, as well as a number of clinical studies that have found that pubertal girls have a higher BMI compared with their similarly aged prepubertal peers.3,31

Longitudinal studies have investigated this issue further. Davison et al9 showed that higher BMI at 5 years and a greater increase in the percentage of body fat between ages 5 and 9 years were associated with earlier puberty in 180 white girls, offering evidence that weight status precedes pubertal timing. However, another longitudinal study of 211 white girls, the Fels Longitudinal

---

**TABLE 1**

Descriptives and Bivariate Comparisons for the Sample According to Puberty Status at Grade 4

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total (n = 354)</th>
<th>Earlier Puberty (Tanner Stage ≥2; n = 168)</th>
<th>Later Puberty (Tanner Stage &lt;2; n = 186)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>9.6 (0.1)</td>
<td>9.6 (0.1)</td>
<td>9.6 (0.1)</td>
<td>.36</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.002</td>
</tr>
<tr>
<td>White</td>
<td>291 (82.2)</td>
<td>127 (75.6)</td>
<td>164 (88.2)</td>
<td>—</td>
</tr>
<tr>
<td>Nonwhite</td>
<td>63 (17.8)</td>
<td>41 (24.4)</td>
<td>22 (11.8)</td>
<td>—</td>
</tr>
<tr>
<td>Maternal education, mean (SD), y</td>
<td>14.6 (2.3)</td>
<td>14.3 (2.2)</td>
<td>14.8 (2.5)</td>
<td>.03</td>
</tr>
<tr>
<td>ITN ratio, mean (SD)</td>
<td>3.7 (3.1)</td>
<td>3.5 (3.2)</td>
<td>3.9 (3.2)</td>
<td>.21</td>
</tr>
<tr>
<td>Age of maternal menarche, mean (SD), y</td>
<td>12.7 (1.5)</td>
<td>12.5 (1.4)</td>
<td>12.9 (1.5)</td>
<td>.01</td>
</tr>
<tr>
<td>Weight status, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal weight (BMI &lt;85th percentile)</td>
<td>257 (72.6)</td>
<td>102 (60.7)</td>
<td>155 (83.3)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>At risk for overweight (85th ≤ BMI &lt; 95th percentile)</td>
<td>52 (14.8)</td>
<td>30 (22.7)</td>
<td>22 (11.4)</td>
<td>.02</td>
</tr>
<tr>
<td>Overweight (BMI ≥ 95th percentile)</td>
<td>45 (12.7)</td>
<td>36 (21.4)</td>
<td>9 (4.8)</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
TABLE 2  BMI, BMI z Score, and Weight Status at Each Age

<table>
<thead>
<tr>
<th>Variable</th>
<th>36 mo (n = 354)</th>
<th>54 mo (n = 354)</th>
<th>Grade 1 (n = 354)</th>
<th>Grade 4 (n = 352)</th>
<th>Grade 5 (n = 297)</th>
<th>Grade 6 (n = 309)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD)</td>
<td>37.4 mo (0.8)</td>
<td>55.9 mo (1.1)</td>
<td>7.0 y (0.3)</td>
<td>9.6 y (0.1)</td>
<td>10.6 y (0.2)</td>
<td>11.6 y (0.1)</td>
</tr>
<tr>
<td>BMI, mean (SD), kg/m²</td>
<td>16.0 (1.3)</td>
<td>16.0 (1.8)</td>
<td>16.7 (2.5)</td>
<td>18.5 (3.8)</td>
<td>19.2 (3.9)</td>
<td>19.9 (4.2)</td>
</tr>
<tr>
<td>BMI z score, mean (SD)</td>
<td>0.18 (0.93)</td>
<td>0.34 (0.98)</td>
<td>0.39 (0.92)</td>
<td>0.39 (1.05)</td>
<td>0.37 (1.05)</td>
<td>0.38 (1.02)</td>
</tr>
<tr>
<td>At risk for overweight, % (n)</td>
<td>14.0 (47)</td>
<td>15.6 (50)</td>
<td>11.4 (36)</td>
<td>16.9 (52)</td>
<td>21.0 (55)</td>
<td>18.0 (49)</td>
</tr>
<tr>
<td>Overweight, % (n)</td>
<td>5.1 (18)</td>
<td>9.3 (33)</td>
<td>10.3 (37)</td>
<td>12.8 (45)</td>
<td>11.8 (35)</td>
<td>11.7 (36)</td>
</tr>
<tr>
<td>Tanner stage ≥2, % (n)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>47.5 (168)</td>
<td>73.3 (214)</td>
<td>92.9 (276)</td>
</tr>
<tr>
<td>Tanner stage ≥3, % (n)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>8.2 (29)</td>
<td>30.8 (90)</td>
<td>56.9 (169)</td>
</tr>
<tr>
<td>Onset of menses, % (n)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.3 (1)</td>
<td>6.3 (21)</td>
<td>23.1 (78)</td>
</tr>
</tbody>
</table>

TABLE 3  Unadjusted Logistic Regression Models Evaluating the Association of BMI z Score at Various Ages Between 36 Months and Grade 6 With Earlier Puberty, Using 4 Different Measures of Puberty

<table>
<thead>
<tr>
<th>Variable</th>
<th>Physical Exam</th>
<th>Maternal Report</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tanner Stage ≥2 at Grade 4 (n = 354, OR (95% CI))</td>
<td>Tanner Stage ≥3 at Grade 5 (n = 292, OR (95% CI))</td>
</tr>
<tr>
<td>BMI z score at 36 mo</td>
<td>1.44 (1.14–1.81)a</td>
<td>1.62 (1.22–2.16)a</td>
</tr>
<tr>
<td>BMI z score at 54 mo</td>
<td>1.45 (1.16–1.82)a</td>
<td>1.53 (1.16–2.03)a</td>
</tr>
<tr>
<td>BMI z score at grade 1</td>
<td>1.98 (1.53–2.56)c</td>
<td>2.03 (1.49–2.75)c</td>
</tr>
<tr>
<td>BMI z score at grade 4</td>
<td>2.26 (1.77–2.87)c</td>
<td>2.22 (1.67–2.96)c</td>
</tr>
<tr>
<td>BMI z score at grade 5</td>
<td>—</td>
<td>2.62 (1.94–3.53)c</td>
</tr>
<tr>
<td>BMI z score at grade 6</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

OR indicates odds ratio; CI, confidence interval. Odds ratios represent the odds of having earlier puberty for each 1-point increase in BMI z score.

a P < .01.
b P < .05.
c P < .001.

TABLE 4  Multiple Logistic Regression Models Predicting the Odds of Having Earlier Puberty Using 4 Different Measures of Puberty

<table>
<thead>
<tr>
<th>Variable</th>
<th>Physical Exam</th>
<th>Maternal Report</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tanner Stage ≥2 at Grade 4 (n = 354, OR (95% CI))</td>
<td>Tanner Stage ≥3 at Grade 5 (n = 292, OR (95% CI))</td>
</tr>
<tr>
<td>BMI z score at 36 mo</td>
<td>1.30 (1.02–1.67)a</td>
<td>1.50 (1.11–2.03)a</td>
</tr>
<tr>
<td>Rate of change of BMI (36 mo to grade 1)</td>
<td>2.47 (1.42–4.30)b</td>
<td>1.62 (0.93–2.80)</td>
</tr>
<tr>
<td>Accelerated BMI</td>
<td>1.30 (0.95–1.77)</td>
<td>1.27 (0.92–1.75)</td>
</tr>
<tr>
<td>Age of mother’s menarche, y</td>
<td>0.88 (0.75–1.03)</td>
<td>0.75 (0.61–0.90)</td>
</tr>
<tr>
<td>Race (other vs white)</td>
<td>2.33 (1.26–4.32)a</td>
<td>1.97 (1.00–3.89)</td>
</tr>
<tr>
<td>Maternal education</td>
<td>0.95 (0.86–1.04)</td>
<td>0.96 (0.85–1.08)</td>
</tr>
</tbody>
</table>

OR indicates odds ratio; CI, confidence interval.

a P < .01.
b P < .05.
c P < .001.

Study, did not find an association between BMI during early childhood and age of menarche. That study only found differences in BMI between early and late maturing girls 4 to 6 years after puberty. Our study and the study by Davison et al had similar proportions of girls classified as at risk for overweight (~16%) and overweight (~13%), comparable to current national estimates of overweight in children, with similar proportions of girls (~50%) reaching Tanner stage 2 puberty at age 9. Although the rates of at risk for overweight and overweight in the population in the Fels Longitudinal Study were not reported, girls were from cohorts spanning the years 1929–1983, well before the increased prevalence of childhood overweight in the United States. Therefore, the lack of differences found in BMI between the early and late maturing groups in the Fels Longitudinal Study may be because of a low proportion of girls with excess weight in that population.

Strengths of our study include the longitudinal study design with BMI data from 36 months of age up to 12 years, the larger sample size, and the socioeconomic and geographic diversity of the sample, which allows for greater generalizability. Furthermore, we were able to adjust for other covariates known to influence the timing of puberty, including age of maternal menarche and race. We found that an earlier age of maternal menarche...
was associated with earlier puberty, which is consistent with studies in the literature.24,34

In addition, although the number of nonwhite subjects in our sample was small, we did find that nonwhite children, of whom most were black, tended to have earlier puberty compared with white children, which has been reported previously in the literature.2,7 We tested for but did not find significant interactions between race and BMI z score in our analysis, which is in contrast to the findings of Kaplowitz et al., who found a stronger relationship between BMI z score and puberty in white girls compared with black girls. It has been hypothesized that the pattern of earlier puberty seen in black girls compared with white girls could be related to the higher rates of at risk for overweight and overweight in black girls.9 The lack of a significant race × BMI interaction in our models would lend support for this hypothesis.

Neither our study nor the study by Davison et al9 used breast palpation for assessing puberty on physical examination, which might lead to misclassification of pubertal status in overweight girls because of fat being confused with breast tissue.20 However our study, as well as the study by Davison et al, used multiple outcome measures of puberty to overcome this limitation. In addition to the standard clinical definition of puberty in girls (Tanner stage 2 puberty or greater breast development) we included the definition of Tanner stage 3 puberty or greater breast development by physical examination, which is a more conservative method for assessing puberty that is less likely prone to misclassification. Furthermore, we also used mother report of Tanner stage 3 puberty or greater breast development, as well as report of onset of menses by a mean age of 12 years, as pubertal measures. Given that mothers were asked to recall the onset of menses as a “yes” or “no” response, this method is less likely prone to recall bias than asking mothers a specific chronologic age at which onset of menses occurred.

Our study had several limitations. Fat mass was assessed with BMI, which is a surrogate measure of adiposity that correlates with fat-free mass, as well as total body fat, and does not account for differences in body fat distribution.35 However, tests such as dual energy radiograph absorptiometry scans for body composition and computed tomography scans for fat distribution are likely not feasible in a large population-based study of girls. Furthermore, in this analysis, we did not evaluate for pubic hair development, which is a sign of increased adrenal androgen production or “adrenarche,” a process that, on average, occurs at about the same time as pubertal breast development but is independent of the hypothalamic-pituitary-gonadal axis.36 Finally, the girls in our study with “earlier” and “later” puberty cannot be considered to have either “precocious” or “delayed” puberty, because we evaluated puberty in a population of girls with normal growth and development starting at a mean age of 9 years.

CONCLUSIONS

Our finding that increased body fatness is associated with earlier pubertal development lends support for the hypothesis that increased rates of obesity among children in the United States may be contributing to a possible secular trend of early maturation in US girls.3,27 Earlier onset of puberty in girls has been associated with a number of adverse outcomes, including psychiatric disorders and deficits in psychosocial functioning17,18, earlier initiation of alcohol use, sexual intercourse and teenage pregnancy39, and increased rates of adult obesity40,41 and reproductive cancers.42 More studies are needed to identify the pathophysiologic mechanisms by which obesity leads to puberty and to determine whether interventions for weight control at an early age may slow or arrest the progression of pubertal onset to earlier ages in the population.

ACKNOWLEDGMENTS

Dr Lee was supported by Pediatric Health Services Research Training Grant T32HD 07534–05 from the National Institute of Child Health and Human Development, National Institutes of Health. This project was supported by the American Heart Association Midwest Affiliate grant-in-aid 0455563Z (to Dr Lumeng).

REFERENCES

10. Zacharias L, Rand WM. Adolescent growth in weight and its


ARTICLE

Pediatric and Emergency Medicine Residents’ Attitudes and Practices for Analgesia and Sedation During Lumbar Puncture in Pediatric Patients

Vicky R. Breakey, MD,<sup>a</sup> Jonathan Pirie, MD<sup>b,c</sup>, Ran D. Goldman, MD<sup>b,c</sup>

Divisions of <sup>a</sup>Hematology and Oncology and <sup>b</sup>Pediatric Emergency Medicine and <sup>c</sup>Pediatric Research in Emergency Therapeutics (PRETx) Program, Department of Pediatrics, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

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

ABSTRACT

OBJECTIVE. Analgesia and sedation for painful procedures in children are safe and effective, yet our experience is that pain management during lumbar puncture is suboptimal. We aim to document factors that influence residents’ decisions to use analgesia and sedation during lumbar puncture and to compare pediatric and emergency medicine residents’ practices.

METHODS. A survey was developed and sent to pediatric and emergency medicine residents from across Canada that inquired about clinical practices, learning experiences, current use of analgesia and sedation for lumbar puncture, and their clinical reasoning for using or abstaining from using analgesia and sedation. The Student’s t and χ² tests were used to compare the 2 resident groups.

RESULTS. Of the 374 residents to whom the survey was sent, 245 completed the survey. Pediatric residents reported performing lumbar punctures with no local anesthetic much more frequently. Pediatric residents used EMLA (AstraZeneca, Wilmington, DE) more frequently and injectable lidocaine less frequently. Pediatric residents used sedation for lumbar puncture at least once, more frequently than emergency medicine residents, and used mostly benzodiazepines. Both groups used ketamine at a similar rate. Pediatric residents reported that they witnessed adverse events of sedation more frequently. Although pediatric residents were responsible for teaching trainees the lumbar-puncture procedure significantly more frequently, they reported less educational opportunities during residency themselves and that they were less likely to recommend the use of local anesthetic during lumbar puncture when teaching the procedure.

CONCLUSIONS. Several significant differences exist between the pediatric residents and emergency medicine residents we surveyed. Pediatric residents were using less injectable local anesthesia for lumbar puncture in children and more sedation for the procedure and have had notably less training in the use of sedation. Pediatric residents have more teaching responsibilities than their emergency medicine
residents colleagues and are inconsistently recommending the use of local anesthetics for lumbar puncture.

LUMBAR PUNCTURE (LP) is one of the most commonly encountered painful procedures in pediatric medicine. Despite evidence to suggest that analgesia and sedation are both efficacious and safe in children, clinical practices do not adhere to these data. Although the LP method is well described in pediatric textbooks, the use of local anesthetics for this procedure remains controversial, and adequate use of analgesia during this painful procedure is limited. A recent report, which used logistic regression analysis, found that LPs performed with local anesthetic by residents and medical students were twice as likely to be successful.

Currently, there are no data on residents’ use of analgesia for painful procedures in children. Procedural skills are taught early in training on an individual basis by a variety of teachers, primarily senior housestaff who are still mastering these skills themselves. Learning to be competent at LP and other procedures is often based on the principle “see one, do one, teach one,” and bad habits may be passed from instructor to learner.

The objectives of our survey were to document and compare the attitudes of pediatric residents (PRs) and emergency medicine residents (ERs) toward the use of analgesics and anesthetics for LPs in children and to describe the current practice as reported by residents across Canada. In addition, we evaluated residents’ education in this area and their LP-teaching methodology.

METHODS
We surveyed residents in accredited training programs in pediatric and emergency medicine across Canada. By contacting the chief residents of each training program, we generated a list of 374 residents from a total of 12 programs. Residents and fellows in pediatric subspecialty programs were excluded because of their variable previous clinical exposure.

We developed a survey that included questions about clinical practices, learning experiences, use of analgesia and sedation for LP, and clinical reasoning for using or abstaining from using analgesia and sedation. We conducted a focus group with fourth-year PRs at the Hospital for Sick Children in Toronto, Canada, to identify key issues. In addition, a pilot study of 10 residents ensured completeness and comprehension of the survey before distribution.

We mailed the survey up to 3 times to each resident between December 1, 2003, and June 1, 2004. To maximize the response rate, envelops were number-coded to determine nonresponders, and second and third surveys were mailed out to nonresponders only. Self-addressed, stamped envelopes were sent to each resident to promote easy return.

Data were collected in Microsoft Excel 2003 (Microsoft Corporation, Redmond, WA). Descriptive analysis of means and SDs for normally distributed continuous variables and skewed continuous data were summarized with medians and interquartile ranges; categorical data were summarized with percentages. The χ² and Student’s t tests were used for comparison of the 2 resident groups by using SPSS 10.0 for Windows (SPSS Inc, Chicago, IL). P values of <.05 were considered significant.

The study was approved by the Hospital for Sick Children’s institutional research ethics board.

RESULTS
Of 374 residents, 245 (67%) completed and returned the survey; 173 (71%) were enrolled in a pediatrics residency, and 72 (29%) were enrolled in either a 5-year Royal College emergency medicine residency program (50) or a 1-year College of Family Physicians of Canada emergency medicine residency program (20). There was representation from all postgraduate years (PGYs) of training: 48 (20%) in PGY1, 59 (24%) in PGY2, 80 (33%) in PGY3, 43 (18%) in PGY4, and 15 (6%) in PGY5 or above. Of the residents surveyed, 231 (94%) had performed at least 1 LP in the preceding 6 months, and 119 (49%) performed between 1 and 4 LPs in the preceding 6 months.

Perception of Pain
Overall, 83 (34%) residents felt that pain was equal among neonates, toddlers, children, and teens undergoing the procedure when asked to estimate the “average pain experienced during LP with no analgesia, anesthesia, or sedation.” The trend, however, showed that the residents surveyed felt the pain during LP to be least in neonates and greatest in children aged 4 to 12 years (Fig 1).

FIGURE 1
Residents’ estimation of the average pain experienced during LP with no analgesia, anesthesia, or sedation on a scale of 1 to 5 (1, painless; 5, excruciating pain) for different age groups.
Local Anesthesia
When asked to describe the frequency of using various methods of local anesthesia for LP, 81 (33%) residents responded that they never perform the procedure without some form of local anesthetic. Comparison of PRs and ERs showed that 57% of PRs frequently, almost always, or always perform the procedure with no local anesthesia versus only 1% of ERs. PRs used EMLA (AstraZeneca, Wilmington, DE) more often than ERs, and the converse was true for lidocaine use (Fig 2).

Factors that discouraged residents from using injectable local anesthesia during LP are presented in Table 1. Age of the patient, pain of injection, and prolongation of procedure were considered significantly more by PRs than by ERs as deterrents to using local anesthetics.

Ten residents (4.1%) reported encountering adverse effects of injectable local anesthesia during LPs. Adverse effects cited included obscured landmarks (3), pain (2), increased anxiety that resulted in increased difficulty of the procedure (2), and local bleeding (1); 2 of these residents did not specify the adverse effect they encountered. There was no significant difference in the frequency of adverse effects from local anesthetic during LPs as reported by PRs and ERs ($P = .386$).

Sedation
Of the residents, 176 (72%) reported that they had used sedation during LPs in children in the past. PRs used sedation more often than the ERs. Of the PRs, 135 (78%) reported frequently, almost always, or always using sedation, compared with 43 (60%) ERs ($P = .004$). Benzodiazepines were the preferred method of sedation, used frequently or more often by 54 (31%) of the PRs and 12 (17%) of the ERs (Fig 3).

Factors that residents reported as discouraging against the use of sedation during LPs are presented in Table 1. Significantly more PRs were deterred from using sedation because of concerns of respiratory depression and the age of the child undergoing the LP than the ERs ($P < .005$).

Of all residents surveyed, 35 (14%) had encountered adverse effects when using sedation. PRs reported more adverse effects than ERs (19% vs 5%; $P = .006$). Adverse effects cited by respondents included paradoxical reaction to midazolam (10), respiratory depression/apnea (9), hypotension (3), prolonged sedation (3), psychosis from propofol (2), patient’s inability to protect the airway (1), increased secretions (1), “ineffectiveness” (1), and seizure (1); 4 residents did not list the adverse effects that they encountered.
Teaching

Of the respondents, 108 (44%) confirmed that they had been educated about the use of sedation in pediatric patients. PRs received less training than ERs (39% vs 57%; \(P = .008\)). Of the residents who reported no training, 114 (85%) suggested that formal training would be useful.

As many as 161 (66%) of those surveyed had taught medical students or other residents the procedure, with a mean of 4.5 trainees per resident-teacher. The PRs reported more teaching responsibilities, with 75% teaching trainees, compared with only 44% of the ERs. Overall, residents reported a significant difference in their teaching recommendations for LP in neonates as compared with LP in children. For neonates, 50% of the residents coach trainees to use local anesthetic, and 12% recommend sedation; for children, 67% teach trainees to use local anesthetic, and 63% recommend sedation.
DISCUSSION
This is the first report of residents’ experiences using analgesia and sedation for LP in children. We chose to compare PRs and ERs on the basis of our hypothesis that training primarily in pediatrics versus adult-based medicine affects decisions in pain management for children. A previous study examined the practices of attending physicians and showed that only 5% (15 of 198) of the pediatric subjects received local anesthesia for LP in a pediatric emergency department, whereas 93% of pediatric patients in a community emergency department were given lidocaine by nonpediatricians. A recent prospective observational study showed that local anesthesia was used in 74% of the infants during LP by medical students and residents and, along with stylet techniques, was associated with a higher LP success rate (defined as cerebrospinal fluid containing <1000 red blood cells per mL).

Our results indicate several significant differences between the PRs and the ERs surveyed. First, it is clear that PRs are using injectable local anesthesia less often for LP in children. Although the use of subcutaneous lidocaine is routine for LP in adults, its use is still debated in children. Lidocaine injections have been shown not to obscure landmarks or hinder the procedure, and the use of injected lidocaine does not reduce the success rate of the procedure, increase the number of attempts, or result in additional trauma. In our study, PRs indicated concern that injected lidocaine is painful, despite evidence that suggests otherwise. Also, PRs cited patient age as a deterrent from the use of local anesthesia, suggesting that neonates’ pain during LP is the most poorly treated of all. However, it has been shown that local lidocaine injection decreases struggling during an LP, which suggests a higher level of comfort for newborns. The use of lidocaine in sick and premature neonates does not increase physiologic instability, and there are no documented disadvantages or adverse effects. One possible reason that ERs are more comfortable using lidocaine is its’ routine use in the adult setting.

PRs, however, use EMLA, a topical anesthetic, more commonly than do ERs for pediatric LPs. EMLA is associated with diminished pain as reported by children during LP and has been shown to be an effective alternative to lidocaine infiltration.

Children can benefit from sedation for painful procedures. Midazolam decreases pain-related and anxiety-related behaviors in children before and after LPs, as observed by both parents and physicians. It also improves pain scores, induces amnesia of the procedure, and decreases anxiety for future procedures. In our study, PRs reported using significantly more sedation for the procedure than did ERs. The reasons for this difference are unclear. It is possible that the fast-paced work in the emergency department decreases the use of sedation for LPs. PRs indicated a significantly higher rate of adverse effects, which may be a result of their lack of training compared with their ER colleagues.

Finally, we found that PRs have more teaching responsibilities than their ER colleagues and are inconsistently recommending the use of injectable local anesthesia for LPs. This raises the concern that procedural training for residents is inadequate. Given the results of this survey, we advocate for the development of more formalized procedural training in pediatric residency programs with training in the use of local anesthetics and sedation for pediatric patients to improve patient care.

There were several limitations to our study. Because no previously validated survey tool exists on this subject, our survey tool was novel. We tried to ensure that the survey was clear and comprehensive by piloting and revising the tool before distribution. Although the response rate was only 67%, it included an adequate sample of residents from all years of training, across the country, and should be generalizable. As with all surveys, there is the possibility of recall bias, and residents may have succumbed to a social-desirability bias and overestimated their attitudes and practices.

CONCLUSIONS
We have documented that incorrect perceptions of pediatric pain persist among residents today. Our results highlight some of the differences that exist in the practices of PRs and ERs. PRs are using less injectable local anesthesia for LP in children, are using more sedation for the procedure, and have less training in the use of sedation. PRs have more teaching responsibilities than their ER colleagues and are inconsistently recommending the use of local anesthetics for LP. The lack of PR education in the use of analgesia and sedation likely contribute to both underuse and misuse, and we recommend pediatric training programs to enhance procedural education to improve care for our pediatric patients.

ACKNOWLEDGMENTS
This study was supported by the Canadian Association of Emergency Physicians Resident Research Grant and the Hospital for Sick Children Trainee Start-up Fund.

REFERENCES
5. Baxter AL, Fisher BL, Burke BL, Goldblatt SS, Isaacman DJ,
How Children With Special Needs Travel With Their Parents: Observed Versus Reported Use of Vehicle Restraints

Taube Korn, MS, OTR\textsuperscript{a,b}, Michal Katz-Leurer, PT, PhD\textsuperscript{a}, Shirley Meyer, MD\textsuperscript{a}, Rosa Gofin, MD, MPH\textsuperscript{b}

\textsuperscript{a}Alyn Hospital, Pediatric and Adolescent Rehabilitation Center, Jerusalem, Israel; \textsuperscript{b}Braun School of Public Health and Community Medicine of the Hebrew University and Hadassah, Jerusalem, Israel

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

ABSTRACT

OBJECTIVES. The difficulties entailed in transporting children with special physical and behavioral needs could influence child restraint misuse and nonuse within this population. Although parental interview is often used to assess child vehicle restraint use, little research had been performed to validate this approach, and none has been done in the special-needs population. The objectives of this study were to assess the prevalence of nonuse and misuse of child restraints in the special-needs population and to assess the validity of using parental report as a measure of child restraint use.

METHODS. Restraint use in 115 children with special needs, aged 0 to 18 years, was observed on their arrival at the parking lot of the Alyn Hospital Pediatric Rehabilitation Center in Jerusalem. The observation noted type of restraint used or absence thereof. If a restraint was used, correct use/misuse was recorded. In 94 cases, the parents were interviewed later that day in the clinic.

RESULTS. Seventy percent of the children were observed as traveling unrestrained or with a restraint that was grossly misused to the extent that it provided no meaningful protection. The remaining children were observed displaying a variety of errors in the selection or use of the restraint that compromised their safety to varying degrees. Analysis of the observation results versus parental reporting revealed a 44% overreporting of child restraint use. Sensitivity was 71%, and specificity was 86%.

CONCLUSIONS. The high prevalence of restraint nonuse and misuse within the special-needs population defines this as a population at risk and emphasizes the need for intervention. Cautious interpretation is required of information acquired from parental reporting of child restraint use. The results of this study should raise awareness among professionals working with children with special needs as to the need for tailored assessment and intervention in the area of child-passenger safety.
A considerable body of research created during the past decade relating to child restraint system (CRS) use within the general population has documented high percentages of child restraint misuse.\textsuperscript{1-3} It can be postulated that the difficulties involved in restraining children with special medical, orthopedic, neuromuscular, and behavioral needs may lead to an even higher proportion of misuse, and possibly nonuse, in this population.

Transporting children with disabilities is complex, and the subject is largely unexplored. Surveys regarding the transportation habits of children with special needs are few,\textsuperscript{4} and some are based on reported behavior. They indicate that these children are at higher risk of injury in case of a crash than are typically developing children. This is because of both the innate physical characteristics of this population, as well as the lack of knowledge among transporters, leading to unsafe modes of transportation. In fact, the lack of information has been found to be one of the underlying reasons for parental concern about the travel conditions of their children with disabilities.\textsuperscript{4-9}

Surveys of reported safety practices are frequently used in injury research, and self-reported safety behavior is often used as an outcome measure in studies that evaluate the effectiveness of prevention interventions.\textsuperscript{10} Well-designed observational studies are time consuming and expensive, whereas interviews, by telephone, in person, or by self-administered questionnaires may be a more efficient way of obtaining information.\textsuperscript{11} However, self-reported safety practices may not accurately estimate safe behavior.

A review of the literature regarding validation of reported vehicle restraint use finds few studies that compare observed child restraint use versus parental reported use and none within a special-needs population. A Canadian study found a 38% parental overreporting of child seatbelt use among observed children entering the parking lot of a children’s hospital whose parents were later interviewed by telephone.\textsuperscript{12} In a similarly designed study in Australia, the observation was conducted as the children were being driven to preschool. This study found a percentage agreement score of 75%, with a low sensitivity of 27% and a high specificity of 99%.\textsuperscript{13} In another study, observed and reported details of car seat use were compared among parents attending car seat checkup clinics in the United States. This study also found a relatively high specificity (>80% in most items), but for nearly every item, a lower sensitivity, ranging from 33% to 74%.\textsuperscript{14} In addition, the literature highlights the phenomenon of increased overreporting in populations with low restraint use.\textsuperscript{11,12,15,16} The objectives of this study were to assess the prevalence of non-use and misuse of child restraints in the special-needs population and to assess the validity of using parental report as a measure of child restraint use in this population.

METHODS

Study Population

The study population included 115 children with special needs. Inclusion criteria for participation in both in-vehicle observation and parental interview were aged 0 to 18 years, attendance in the multidisciplinary outpatient clinics of the Alyn Hospital Pediatric and Adolescent Rehabilitation Center in Jerusalem, and participation in the clinic’s occupational therapy (OT) evaluation. Exclusion criteria included children who arrived via public transportation and children who arrived accompanied by someone other than a parent (a parent will, from here on, be referred to in the masculine regardless of whether the mother or father was the study participant).

Ninety-four children met all of the inclusion criteria. An additional 21 children were observed in their vehicles but did not meet the third inclusion criteria of participation in the OT evaluation on the day of the clinic, and, thus, the parents of these children could not be interviewed. These 21 children were only included in the part of the study evaluating prevalence of restraint use and misuse. No significant difference was found between diagnostic categories or sociodemographic characteristics (age, ethnic origin, religion, and parental age and level of education) of the children who only participated in the observation as compared with the children whose parents were also interviewed. Sample size was calculated to detect a difference in reported versus observed behavior of 40%\textsuperscript{12} with an \( \alpha \) value of .05 and a power of .80.

The response rate of parents approached for observation of their child was 96%. Among those parents who did not agree to participate in the study, the reason usually given was that they were late for their clinic appointment. One-hundred percent of the parents who were asked to be interviewed, agreed. This study was approved by the hospital’s ethics committee.

Data Collection

Observations were conducted by occupational therapists who were trained as child-passenger safety (CPS) technicians or technician instructors.\textsuperscript{1} The observations were conducted between June 2004 and February 2005 on mornings when multidisciplinary outpatient clinics were held. The observer received a list of children scheduled for the clinic on the given day. All families were stopped as their car arrived at the rehabilitation center’s parking lot, and the observer determined whether the child in the vehicle was on the clinic list. If so, the parent was given a brief explanation regarding the study and was asked to sign an informed consent form. In consideration of the center’s multiethnic population, translation services were provided as needed.

The observation tool was based on checklists used by
CPS technicians, providing face and consensual validity. This form included information on the child’s age and weight, seating position in the car, use or nonuse of a restraint, and characteristics of the vehicle that the child arrived in. In addition, it listed the different components of child restraint use, noting correct use or misuse of each of the components. Information on the use of orthopedic or medical equipment, travel while seated in a wheelchair, seat belt use, and lack of restraint use was included. Correct use was defined as use of the restraint according to the manufacturer’s instructions and National Highway Traffic Safety Administration curriculum instructions. Misuse was defined as any deviation from these instructions. Gross misuse was defined as a child sitting in a CRS not anchored to the vehicle’s seat by a seat belt and/or a child not restrained by the internal harness of the CRS system, a child sitting in a wheelchair anchored to the vehicle and the child not restrained by the passenger portion of the restraint, or a restrained child sitting in a wheelchair not anchored to the vehicle. Interrater reliability for observation, based on a comparison of the ratings of 5 children by 2 observers, was 100%.

Reported data were collected in the form of parental interviews during the OT evaluation conducted as part of the multidisciplinary clinic. The occupational therapist in the clinic was blind to the observer’s earlier data collection. Data regarding reported behavior were collected using a structured interview based on a closed questionnaire that mirrored the items noted by the observer. Demographic information was added to this form, as well as information regarding distance traveled to the center and information regarding the child’s medical and orthopedic condition. If the family indicated restraint nonuse, an open question was added inquiring as to the reason for nonuse. Interrater reliability for interview, based on a comparison of the ratings of 4 parents by 2 interviewers, was 95%.

After the survey, participants received a letter including CPS information and an invitation to participate in individual hands-on CPS instruction. When possible, this instruction was provided on the day of the clinic visit after the completion of the observation and interview.

Data Analysis

Observed results were selected as the best available measure of restraint use for the study of validity. Sensitivity was defined as the proportion of unrestrained children who are reported by their parents as unrestrained. Specificity was defined as the proportion of restrained children who are reported by their parents as restrained. Positive predictive value (PPV) was defined as the likelihood that the child is not restrained when the parent describes him as not restrained. Negative predictive value (NPV) was defined as the likelihood that the child is restrained when the parent describes him as unrestrained. The percentage of overreporting is the likelihood that the parent describes their child as restrained when they are, in fact, not restrained. Percentage of agreement was defined as those whose positive and negative answers coincide, as a percentage of the total population. In all of the calculations, gross misuse and nonuse were reported as 1 category, as gross misuse indicates a level of protection equivalent to nonuse. For all of the calculations, level of significance was defined as P < .05. Analyses were conducted using SPSS.

RESULTS

Characteristics of the Study Population

Sociodemographic characteristics of the study population are presented in Tables 1 and 2. There were more male than female children, with more than half <10 years of age. Close to half of the parents had not completed 12 years of education, and slightly more than half the population was Arab.

Thirty percent of the children in the study population were diagnosed with spina bifida, 37% with traumatic brain injury or cerebral palsy, and 23% with neuromuscular disorders. Just more than half of the children weighed ≤18 kg. Ten children were observed traveling while wearing a reciprocal gate orthosis, 8 children were wearing a thoracic lumbar sacral orthosis, and 2 were transported with an oxygen supply unit.

Restraint Use/Misuse

Half (50%) of the children were observed in the vehicle with no restraint at all. This included 56 children who were unrestrained on the vehicle seat and 2 who were unrestrained in wheelchairs (Table 3). Most (89%) of the children observed were sitting in the back seat, and the remainder were sitting beside the driver.

The inclusion of children in a grossly misused CRS in the “no restraint” category (as explained in the definitions of variables) brings the number from 58 to 80 children (~70%) in this study who were observed as

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Sociodemographic Characteristics of the Study Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>n</td>
</tr>
<tr>
<td>Child’s gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>69</td>
</tr>
<tr>
<td>Female</td>
<td>46</td>
</tr>
<tr>
<td>Total</td>
<td>115</td>
</tr>
<tr>
<td>Child’s age, y</td>
<td></td>
</tr>
<tr>
<td>0–4</td>
<td>37</td>
</tr>
<tr>
<td>5–9</td>
<td>34</td>
</tr>
<tr>
<td>≥10</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>115</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Jewish</td>
<td>55</td>
</tr>
<tr>
<td>Arab</td>
<td>60</td>
</tr>
<tr>
<td>Total</td>
<td>115</td>
</tr>
</tbody>
</table>
traveling without restraint protection. The gross misuse occurred both among the CRS users and wheelchair tie-down system users. Of the 35 children (30% of the total study population) who were restrained, all displayed misuse either in choice or use of the restraint, compromising their safety to varying degrees.

Report Versus Observation
The proportion of children accurately reported as restrained (specificity) was higher than the proportion of children accurately reported as unrestrained (sensitivity). Slightly more than 50% of the time, a parental report of restraint use was an indication that a restraint was indeed used (NPV). Most of the time, when a parent said that he did not use a child restraint, this was indeed the case (PPV). Forty-four percent of parents overreported restraint use (Table 4).

Reported Reasons for Nonuse of Child Restraints
Almost one third (32%) of the parents who provided reasons for reported nonuse of child restraints described physical and behavioral needs that precluded use. Other reasons given were less related to the child’s special needs but rather reflected parental knowledge, beliefs, and economic barriers to restraint use (Table 5).

DISCUSSION
Use/Misuse
The high percentages of lack of restraint use and restraint misuse found in this study indicate that almost all of the children surveyed would be at high risk for injury in a motor vehicle crash. The identification of children with special needs as a group at risk for injury in motor vehicle crashes is important for public health action aimed at designing and implementing intervention strategies for this group. As emphasized by parents who commented on why they did not restrain their children, children with special physical and behavioral needs provide a challenge even for safety-minded parents who wish to restrain their children but who do not always have the proper tools. Indeed, several cases were observed of parents, particularly of older children, who improvised a solution to provide their child with a means of sitting in the car. These solutions were not safe but were the best that the parent could do with the information they had.

Another factor that may have contributed to the high proportion of nonuse found in this study is the low socioeconomic status (SES) of the study population. SES, often measured by education, income, and/or occupation, has, in general, been found to be positively associated with safety behaviors and, in particular, with restraint use. Possible explanations for the direct relationship between SES and safety behaviors and the in-
verse relationship between SES and childhood injury include the influence of low parental education on safety knowledge. In addition, the cost involved in purchasing safety devices may serve as a deterrent for low-income families.10,21

The percentage of misuse found in this study is in line with the literature surveying the general population. Studies using similarly detailed and sensitive tools to those used in this study, at car seat checks, also find high rates of misuse, despite the high awareness about the importance of CRS use in those study populations.14,22,23 Two studies focusing on low SES populations found proportions of nonuse that were even greater that those found in the present study.24,25

Parental Report Versus Observation

The present study found 44% parental overreporting of child restraint use. This, as well as the percentage agreement score, is similar to findings in studies in the general population.12,13 In the present study, as in the literature, sensitivity was lower than specificity. The high PPV indicates that a parental report of lack of child restraint use may be relied on as accurate, whereas a report of child restraint use may not be accurate.13,14

In light of the literature regarding increased overreporting where there is low restraint use, the high level of overreporting found in this study may be expected considering the low prevalence of restraint use. This phenomenon may be because of the tendency for people to often report restraint use whether true or not. When the actual prevalence of restraint use is low, there will be more people available to give a false answer, and, therefore, the proportion of overreporting is higher.

Why do parents report something that is not true? One reason cited in the literature is the tendency to give a socially desirable response, that is, a response in accordance with what a good parent would do.12,13,15,16 In the present study setting, parents were interviewed by a professional with the status of an authority figure from whom they receive care. The parent may want to seek approval from the authority figure and be seen as a “good parent.” Alternatively, the parent may want to finish the interview quickly without being delayed by an explanation of why he should be restraining his child. He may feel that he knows the official line on the topic but that he is not personally convinced of the importance of child restraint use in its absolute sense or in the context of all of the other tasks required in caring for a child with special needs. Another plausible explanation is that the parent is answering the question in relation to how he knows he should transport his child or would like to transport his child or in relation to how he sometimes transports his child. As pointed out in the literature, parental recall of events is reinforced, among other things, by their view of the world and how they think things should be.26

There was a small percentage of parents whose children were observed as restrained but who reported lack of restraint use. The lack of coordination between the parent responsible for the child’s care in the car and the parent who answered the interview questions may provide us with an explanation for this finding.

It is possible that parental knowledge of observation of their child in the vehicle could have biased the results of the subsequent interview. Although foreseen as a possible study limitation, the observation/report sequence used in this study seemed to be the best alternative, because observation as the child arrived in the morning gave the most accurate indication of restraint use. There is some additional evidence in the literature regarding the existence of overreporting even when the individual knows that his behavior has been observed.11,27 If a bias does exist in this study, then an even higher degree of overreporting exists in reality as compared with the high percentage found in this study, emphasizing the inaccuracy of using parental reporting as a measure of child restraint use.

This study has some limitations. The study population was confined to children treated at 1 center. Israel is a small country with few centers providing the services provided at the study center. It is estimated that about 70% of children in Israel with complex disabilities resulting from a variety of diagnoses are treated at this center, allowing the results to be generalized. Nevertheless, additional research is needed to confirm the ability to generalize from the findings to other populations.

CONCLUSIONS

The present study provides important baseline information regarding restraint use and misuse for children with special needs. The high prevalence of restraint nonuse and misuse in the special-needs population defines this as a population at risk and emphasizes the need for tailored intervention. The barriers to restraint use and correct use that are particular to this population must be taken into account when planning effective intervention.

The high percentage of parental overreporting of child restraint use requires us to cautiously interpret information acquired by parental interview and points to the limitations of using parental reporting as an indicator of individual child restraint use and as an estimate of population restraint use for policy planning. It highlights the need for objective observational evaluation, despite this being costly and time consuming.

ACKNOWLEDGMENTS

We thank the administration and staff of the Alyn Hospital-Pediatric and Adolescent Rehabilitation Center for their support of this work and, particularly, the Occupational Therapy Department for its assistance in data collection.
REFERENCES

26. Scott I. You can’t believe all that you are told: the issue of unvalidated questionnaires. Inj Prev. 1997;3:5–6
Admission Temperature of Low Birth Weight Infants: Predictors and Associated Morbidities

Abbot R. Laptook, MD; Walid Salhab, MD; Brinda Bhaskar, MS; and the Neonatal Research Network

ABSTRACT

BACKGROUND. There is a paucity of information on the maintenance of body temperature at birth for low birth weight infants.

OBJECTIVES. We examined the distribution of temperatures in low birth weight infants on admission to the NICUs in the Neonatal Research Network centers and determined whether admission temperature was associated with antepartum and birth variables and selected morbidities and mortality.

METHODS. Infants without major congenital anomalies born during 2002 and 2003 with birth weights of 401 to 1499 g who were admitted directly from the delivery room to the NICU were included. Bivariate associations between antepartum/birth variables and admission temperature and selected morbidities/mortality and admission temperature were examined, followed by multivariable linear or logistic regressions to detect independent associations.

RESULTS. There were 5277 study infants and the mean (±SD) birth weight and gestational age were 1036 ± 286 g and 28 ± 3 weeks, respectively. The distribution of admission temperatures was 14.3% at <35°C, 32.6% between 35 and 35.9°C, 42.3% between 36 and 36.9°C, and 10.8% at ≥37°C. The estimate of birth weight on admission temperature with and without intubation was +0.13°C and +0.04°C per 100-g increase in birth weight, respectively. The mean admission temperature for each center varied from 1.5°C below to 0.3°C above a reference center. On adjusted analyses, admission temperature was inversely related to mortality (28% increase per 1°C decrease) and late-onset sepsis (11% increase per 1°C decrease) but not to intraventricular hemorrhage, necrotizing enterocolitis, or duration of conventional ventilation.

CONCLUSIONS. Preventing decreases in temperature at birth among low birth weight infants remains a challenge. Associations with intubation and center of birth suggest that assessment of temperature control for infants intubated in the delivery room may be beneficial. Whether the admission temperature is part of the causal path or a marker of mortality needs additional study.
Birth is associated with changes that affect the body temperature of the newborn. These include the ambient room temperature, multiple routes of heat loss (evaporative, convective, and conductive), and increases in oxygen consumption with consequent heat production. Heat loss usually far exceeds heat production after birth, and if measures are not initiated to reduce heat loss, body temperature will fall. An excessive fall in body temperature may impair the transition from intrauterine to extraterine circulatory pathways given the effect of temperature on pulmonary vasomotor tone and acid-base homeostasis. In general, effective interventions to prevent cold stress for the term infant are applied to the preterm infant, for example, drying and the use of radiant warmers; however, a higher surface area/weight ratio and skin characteristics make reducing heat loss for the preterm infant a more formidable challenge. Current data on the relative success or failure in avoiding cold stress for preterm newborns has been limited to the extremes of prematurity or small numbers of patients from third-world countries. The purpose of this report was to use a large multicenter cohort of low birth weight infants to determine the following: (a) the frequency distribution of temperatures on admission to NICUs, (b) the variables at birth that are associated with the largest extent of reduced admission temperature, and (c) whether admission temperature is independently associated with selected neonatal morbidities and in-hospital mortality.

METHODS

The study was conducted among 15 centers of the National Institute of Child Health and Human Development Neonatal Research Network. Data were retrieved on all of the neonates with the following inclusion criteria: born at a network center between January 1, 2002, and December 31, 2003; born with a birth weight of 401 to 1499 g; admitted directly to a NICU from a delivery room; and born without a major congenital anomaly. The first temperature obtained on admission of each infant to the NICU from the labor and delivery department was recorded as the admission temperature along with the date and time charted. Temperatures recorded in the delivery room or during transport to the NICU were not recorded. The site of temperature measurement (axilla, skin, or rectal) was noted. Exclusion criteria were missing admission temperature, missing time of admission temperature, or temperature recorded after 2 hours of age.

Data on each mother and infant were prospectively collected as part of an ongoing survey of neonatal morbidity and mortality initiated in 1987. Trained research nurses reviewed the medical charts of mother and infant and entered predefined data items into an institutional review board-approved computerized database. Neonatal outcome data were assessed at discharge from the hospital, 120 days after birth, or at the time of death, using which ever came first.

Variables explored for associations with admission temperature included the following: (a) maternal variables, including exposure to antibiotics (any use during the hospitalization for delivery), tocolytics, antenatal steroids (partial or complete course), and the presence of multiple births; (b) intrapartum variables, including the presence of labor, ruptured membranes >18 hours, and the mode of delivery; (c) infant characteristics, including birth weight, gestational age (obstetric criteria), and gender; (d) delivery room variables, including intubation and/or chest compression with or without resuscitative medications, Apgar scores, and umbilical artery pH and base excess; (e) site (axilla, rectum, or skin) and age of temperature measurements; and (f) network center of birth. Neonatal outcomes included days of conventional ventilation, late-onset sepsis (positive blood culture after 72 hours of age), necrotizing enterocolitis (NEC; modified Bell’s stage IIa or above), grade III or IV intraventricular hemorrhage (IVH), and death after 12 hours of age and before hospital discharge. The assigned cause of death reflects the purported underlying, proximate disease process contributing to death and is based on autopsy and clinical findings using predefined causes in the manual of operations for the database.

Data analysis for associations with admission temperature were initially explored with bivariate analyses between admission temperature and maternal and intrapartum variables, infant characteristics, and delivery room events. Variables significant at a .10 level of significance in bivariate analyses were entered into multivariable linear regressions. Umbilical artery pH and base deficit were not included in the multivariable analysis, because values were available only for a subset of the cohort. Gestational age and chest compressions/resuscitative medications were not included because of colinearity with birth weight and intubation, respectively. One center was designated the reference center on the basis of the highest percentage of admission temperatures between 36 and 36.9°C (center 10). Center results in multivariable analyses were expressed relative to center 10.

In a similar fashion, analyses for associations between admission temperature and outcomes were initially explored with bivariate analyses, and variables significant at a .10 level of significance were entered into multivariable linear regressions for continuous outcomes and logistic regressions for categorical outcomes. These analyses were controlled for antenatal steroids, gender, race, birth weight, intubation, Apgar at 5 minutes, and center. Results of logistic regressions were expressed using odds ratios (ORs) and 95% confidence intervals (CIs). Results of multivariable linear regressions were expressed using the parameter estimate to indicate the magnitude of independent associations.
RESULTS

Between January 1, 2002, and December 31, 2003, there were 7498 infants entered into the database. Applying the inclusion criteria resulted in the exclusion of 1649 infants, of which 1450 were excluded for 1 criteria only (749 outborn, 462 not directly admitted to an NICU, 207 with an anomaly, and 32 out of the weight range). Applying the admission temperature criteria to the remaining 5849 infants resulted in the exclusion of 570 infants. Of the latter infants, 264 were considered viable (given delivery room interventions, ventilator support, intravenous fluids, etc), and 306 were nonviable (no care provided). All but 59 of the 570 infants excluded for the temperature requirement had missing temperature or time of temperature. Two additional infants fulfilled the temperature requirement but were nonviable and were excluded. The study cohort was composed of 5277 infants.

Descriptive characteristics of selected maternal and intrapartum variables, infant characteristics, and variables from the delivery room are listed in Table 1. Infants excluded for not meeting the admission temperature criteria (n = 570) were of a lower birth weight (880 ± 318 and 545 ± 151 g, mean ± SD for viable and nonviable, respectively) and gestational age (26.3 ± 3.2 and 22.7 ± 2.3 weeks for viable and nonviable, respectively). Umbilical artery blood gas results were available for 48% (pH data) and 46% (base deficit data) of the cohort. The mean admission temperature was 35.9 ± 1.0°C (range: 28–39.6°C). The distribution of admission temperatures among the cohort (Fig 1) demonstrates that 46.9% of the temperatures were <36°C. In contrast, the frequency of admission temperatures ≥37.0°C was 10.8% and ≥38°C was 1.3%. The frequency of admission temperatures <35 and <36°C increased with decreasing gestational age and birth weight (Table 2). The measurement site of admission temperatures varied on the basis of center practice and was recorded from the axilla (77.6%), rectum (15.5%), and skin (7.0%); 9 patients had missing data for this item. The mean age at the admission temperature was 23 ± 14 minutes with a median value of 20 minutes (25th and 75th percentiles of 14 and 27 minutes, respectively).

Variables present before or at birth and associated with admission temperature on bivariate analyses were multiple births, labor, use of antenatal steroids, maternal antibiotics, rupture of membranes, mode of delivery, birth weight, gestational age, center, intubation, 5-minute Apgar, site and age of temperature measurement, and center. Only the variables listed in Table 3 were significantly associated with the admission temperature on multivariable analyses. Multiple births, use of antenatal steroids, and prolonged rupture of membranes were associated with a statistically significant but small change in admission temperature, each <0.2°C compared with the absence of the variable. A similar change in temperature was associated with age of temperature measurement. In contrast, birth weight and intubation, center of birth, and the Apgar score at 5 minutes were associated with the largest change in admission temperature (Table 3). There was a significant interaction between birth weight and intubation in the delivery room. The admission temperature was 0.04°C higher with each 100-g increase in birth weight; however, for infants requiring intubation in the delivery room, the admission temperature was 0.13°C higher with each 100-g increase in birth weight. The admission temperature was 0.05°C higher for each point increase in the Apgar score at 5 minutes. The site of temperature measurement was associated with the admission temperature in that rectal and axilla temperatures were 0.40 and 0.22°C higher than skin temperature, respectively. Finally, there was a prominent association between center of birth and the admission temperature. The average admission temperature of each of the 14 centers ranged from 0.3°C above to as much as 1.5°C below the average admission temperature of the reference center. The variability in the distribution of admission temperatures among the 15 centers is plotted in Fig 2.

The frequencies of selected neonatal morbidities for this cohort were 6.3% for NEC, 10.3% for IVH grades III and IV, 23.3% for late-onset sepsis, and 10 ± 18 days of conventional ventilation. In-hospital mortality was 12.2% with 45.2% of the deaths occurring at <7 days of age. Major categories of assigned causes of death are listed in Table 4. In multivariable analyses, there was no association between the admission temperature and NEC (OR: 1.0; CI: 0.90–1.16), grade III/IV IVH (OR: 0.96; CI: 0.86–1.07), or duration of conventional ventilation (0.4 days per 1°C decrease in the admission temper-
In contrast, for every 1°C decrease in admission temperature, the odds of late-onset sepsis were increased by 11% (OR: 1.11; CI: 1.02–1.20), and the odds of dying were increased by 28% (OR: 1.28; CI: 1.16–1.41).

**DISCUSSION**

This report documents the temperature on admission of a recent large cohort of very low birth weight infants born within 15 academic centers and transferred directly to the NICU from the labor and delivery department. The principal findings of this study are as follows: (a) low temperatures on admission are common; (b) there are important associations between the admission temperature and variables antecedent to admission that may be amenable to change; and (c) there is a prominent association between the extent of reduced temperature on admission and both late-onset sepsis and in-hospital mortality.

Efforts to limit heat loss are important initial steps in the stabilization of newborns immediately after birth and are incorporated in the Neonatal Resuscitation Program and the World Health Organization’s guide to thermal control of the newborn. Minimizing heat loss in low birth weight and premature infants is difficult because of high evaporative heat loss exacerbated by a large temperature gradient from the skin to the ambient air and physical characteristics of the premature infant (increased surface area/weight ratio, immature epidermal barrier, limited vernix caseosa, and subcutaneous fat). There are relatively few reports on the frequency of low temperatures at birth among premature infants. Among hospitals in third-world nations, the frequency of temperatures at 2 hours after birth can be as high as 60% for cohorts that include both term and low birth weight infants, and these observations are linked to a high incidence of hypothermia at 24 hours of age. Even in developed countries, the frequency of admission temperature <35°C at 2 hours after birth can be as high as 60% for cohorts that include both term and low birth weight infants, and these observations are linked to a high incidence of hypothermia at 24 hours of age.

![Figure 1](https://example.com/figure1.png)

**Figure 1**

Results are presented for the distribution of admission temperatures among 5277 low birth weight infants irrespective of measurement site (axilla, rectum, or skin).

<table>
<thead>
<tr>
<th>Gestational Age, wk</th>
<th>n</th>
<th>Birth Weight, mean ± SD, g</th>
<th>Admission Temperature, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>&lt;35°C</td>
</tr>
<tr>
<td>28</td>
<td>643</td>
<td>1088 ± 201</td>
<td>9.6</td>
</tr>
<tr>
<td>27</td>
<td>609</td>
<td>977 ± 182</td>
<td>10.7</td>
</tr>
<tr>
<td>26</td>
<td>539</td>
<td>840 ± 163</td>
<td>13.2</td>
</tr>
<tr>
<td>25</td>
<td>468</td>
<td>751 ± 130</td>
<td>20.5</td>
</tr>
<tr>
<td>24</td>
<td>397</td>
<td>655 ± 100</td>
<td>33.8</td>
</tr>
<tr>
<td>&lt;24</td>
<td>187</td>
<td>598 ± 118</td>
<td>43.9</td>
</tr>
</tbody>
</table>

Data are presented up to 28 weeks, because the registry is defined by birth weight, and infants >28 weeks with a birth weight >1500 g will not be included.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter Estimate, °C</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple births</td>
<td>0.05</td>
<td>.045</td>
</tr>
<tr>
<td>Use of antenatal steroids</td>
<td>0.10</td>
<td>.006</td>
</tr>
<tr>
<td>Rupture of membranes &gt;18 h</td>
<td>0.19</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Mean age at admission temperature</td>
<td>0.17</td>
<td>.0023</td>
</tr>
<tr>
<td>Intubation × birth weight&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight with intubation, per 100-g increase</td>
<td>0.13</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Birth weight without intubation, per 100-g increase</td>
<td>0.04</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>5-min Apgar score, per Apgar point increase</td>
<td>0.05</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Center: lowest/highest average value&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-1.51/0.29</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Measurement site: rectal/axillad&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.40/0.22</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

<sup>a</sup> The parameter estimate indicates the magnitude of change in temperature (°C) in the presence of the variable listed. The overall model had an r² value of 0.33 (P < .0001).

<sup>b</sup> There was a significant interaction between birth weight and intubation that resulted in different parameter estimates for birth weight on the basis of the presence or absence of intubation.

<sup>c</sup> Values for each center are relative to the reference center. For simplicity, the centers with the lowest and highest average admission temperature are listed.

<sup>d</sup> Values are relative to skin temperature.

The temperature; P = .1). In contrast, for every 1°C decrease in admission temperature, the odds of late-onset sepsis were increased by 11% (OR: 1.11; CI: 1.02–1.20), and the odds of dying were increased by 28% (OR: 1.28; CI: 1.16–1.41).
basis of obstetric criteria; mean of 28 weeks) defined by birth weight (<1499 g) and may account for the lower frequency of admission temperatures <35°C (14.3%). However, in this cohort, ~47% of the admission temperatures were <36°C. An important limitation of this study is the observational design without a standard practice regarding site, time, device, and technique used for temperature measurements. In addition, no data are available on the maternal temperature at the time of or immediately proximate to delivery, the temperature of the delivery room, or the qualifications of the pediatric providers in attendance.

Interventions to minimize the extent of heat loss have been studied in small groups of infants. The most effective seems to be occlusive wraps for which there has been interest over the past 30 years.21 More recently, 3 randomized clinical trials demonstrated that the use of polyethylene wraps or polyurethane bags compared with drying in the delivery room prevented heat loss and better maintained rectal admission temperatures for infants <29 weeks' gestation.22–24 Additional measures include the use of caps, which have been demonstrated to reduce the exchange of heat between the head and the ambient air.25,26 The use of caps conveniently complements the application of occlusive wraps where heat loss from the exposed head is still a concern. Important questions regarding the use of occlusive wraps are whether there are low-frequency adverse effects on the skin, alterations in skin flora, or potential overheating of the body. Systematic reviews and formal meta-analysis suggest that elevated temperatures recorded on NICU admission among wrapped infants may reflect factors such as maternal temperature and infection rather than occlusive wraps and indicate the need for additional studies.27,28 Information regarding the use of occlusive wraps or other means to reduce heat loss from the current cohort was not collected as part of this study. Given the frequency of low admission temperatures among infants in this report, surveillance of temperatures in the delivery room and on admission would seem to be an appropriate, worthwhile, quality improvement initiative.

Variables antecedent to and independently associated with prominent changes in admission temperature were birth weight and intubation, Apgar score at 5 minutes, and center of birth. The association with birth weight was expected, because the physical characteristics of low birth weight infants predispose to a mismatch between heat production and heat exchange with the ambient environment when high-risk infants are stabilized at birth.14 This is consistent with the EPICure Study, where...

---

**TABLE 4** Causes of Death

<table>
<thead>
<tr>
<th>Causea</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiratory distress</td>
<td></td>
</tr>
<tr>
<td>Isolated</td>
<td>88</td>
</tr>
<tr>
<td>With severe intracranial hemorhage</td>
<td>71</td>
</tr>
<tr>
<td>With infection</td>
<td>49</td>
</tr>
<tr>
<td>NEC</td>
<td>47</td>
</tr>
<tr>
<td>with sepsis</td>
<td>42</td>
</tr>
<tr>
<td>Sepsis (early and late onset)</td>
<td>70</td>
</tr>
<tr>
<td>Bronchopulmonary dysplasia</td>
<td></td>
</tr>
<tr>
<td>Isolated</td>
<td>26</td>
</tr>
<tr>
<td>With infection</td>
<td>29</td>
</tr>
<tr>
<td>With brain injury</td>
<td>10</td>
</tr>
<tr>
<td>Severe intracranial hemorhage</td>
<td>16</td>
</tr>
<tr>
<td>Immaturity</td>
<td>46</td>
</tr>
</tbody>
</table>

These causes represent 77% of the 642 deaths in this cohort of 5277 infants. The remainder of the assigned causes reflects multiple conditions or suspected diagnoses.
an admission temperature <35°C occurred in 30%, 43%, and 58% of infants of 25, 24, and 23 weeks' gestation (median birth weights of 760, 680, and 600 g, respectively). Multivariate analysis in the present study indicates that the effect of birth weight alone is relatively small, with an average of a 0.4°C difference in admission temperature between infants with birth weights of 401 and 1499 g. In contrast, if intubation is performed, the effect of birth weight is an average of 1.4°C different between infants of the same 2 birth weights. The association of the Apgar score at 5 minutes may parallel the birth weight-intubation interaction, because lower Apgar scores may represent a proxy for infants having more resuscitative measures (of which the most common procedure is intubation) or the response to such interventions. Center of birth was associated with a prominent change in admission temperature relative to the reference center. Five of the 14 centers had an average admission temperature >0.75°C lower than the reference center, and 3 centers were >1°C lower. Delineation of the specific practices for maintenance of temperature in the delivery room at each center was not part of this study, but the results suggest that a quality improvement approach using benchmark initiatives may be helpful for some centers.

The admission temperature was not associated with NEC, severe IVH, or duration of conventional ventilation. Associations were present between admission temperature and both late-onset sepsis and in-hospital mortality. Thermal management has been labeled a cornerstone of neonatology. The latter is based on the pioneering work of Silverman et al that maintenance of body temperature through control of the thermal environment during the first 5 days of life (isoelette temperature of 29 vs 32°C with resultant axillary temperatures of 31.1 vs 33.7°C, respectively) reduced mortality in low birth weight infants. Other clinical trials of low birth weight infants yielded similar observations. These therapeutic trials outlined the effects of a thermal management scheme on mortality rather than an association between admission temperature and outcome.

Additional temperatures beyond admission to the NICU were not collected on infants in the present cohort. Whether prevention of low temperatures at birth decreases mortality or whether the low admission temperature is part of the casual path or simply a marker for an increase in the odds of mortality cannot be determined from this observational analysis. Previous investigations that have reported associations between admission temperature and mortality have insufficient sample size, were not adjusted for covariates, and were not reproducible. The association in this report between admission temperature and late-onset sepsis provides a potential path to link the admission temperature and mortality. In adults, a self-limited interval of perioperative hypothermia may promote postoperative infections via temperature-mediated impaired immune function; perioperative normothermia decreased the postoperative infectious complication. Whether late-onset sepsis remote from birth is causally linked to admission temperature is unknown. In addition, the causes of death (Table 4) seem to reflect expected complications of prematurity.

The results of this observational cohort demonstrate that minimizing the extent of temperature reduction at birth for the low birth weight and premature infant remains challenging. The birth weight-intubation interaction with the admission temperature and the variability among the various participating centers suggest that thermal control for newborns requiring respiratory support at birth requires a reassessment of practice. The time, effort, and resources to determine whether avoidance of temperature reductions at birth reduces mortality seem to be well justified in view of a potential casual path via late-onset sepsis.

REFERENCES


Lead Exposure, IQ, and Behavior in Urban 5- to 7-Year-Olds: Does Lead Affect Behavior Only by Lowering IQ?

Aimin Chen, MD, PhD, Bo Cai, PhD, Kim N. Dietrich, PhD, Jerilynn Radcliffe, PhD, Walter J. Rogan, MD

Epidemiology Branch and Biostatistics Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina; Department of Environmental Health, University of Cincinnati, Cincinnati, Ohio; Department of Psychology, Children’s Hospital of Philadelphia and University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania

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

ABSTRACT

BACKGROUND. Lead exposure in childhood lowers IQ scores, but its effect on children’s behavior is less clear. Because IQ, per se, affects behavior, measuring the direct effect of lead requires measuring and then adjusting for IQ. In addition, either peak blood lead concentration, usually at 2 years old, or the lower blood lead level measured at school age may be the most relevant. Few studies have all of this information.

OBJECTIVES. The purpose of this work was to differentiate the direct effect of lead on behavior and the indirect effect through IQ and to examine the strength of the association for peak and concurrent blood lead concentration.

METHODS. Data come from a clinical trial of the chelating drug succimer to prevent cognitive impairment in 780 urban 12- to 33-month-olds with blood lead concentrations of 20 to 44 μg/dL. The children were followed from ages 2 to 7 years. The trial data were analyzed as a prospective observational study.

RESULTS. Blood lead concentration at 2 years old was not associated with Conners’ Parent Rating Scale-Revised scores at 5 years of age or Behavioral Assessment Systems for Children scores at 7 years of age. Blood lead level at 7 years of age had direct effects on the Behavioral Assessment Systems for Children behavioral symptoms index, externalizing, and school problems at age 7.

CONCLUSIONS. Concurrent blood lead concentration was associated with externalizing and school problems scales at 7 years of age, and the effect was not entirely mediated through the effect of lead on IQ.
Even small amounts of lead exposure in childhood seem to lower scores on IQ tests. Multiple cohort and cross-sectional studies give similar estimates of the size of the effect.1–5 Although the Centers for Disease Control and Prevention had set a “level of concern” at 10 μg/dL blood lead concentration, a threshold below which lead does not affect IQ has not been determined, and recent studies have extended the relation to <10 μg/dL.6–8 Whether lead exposure produces other psychological or behavioral damage in children is less well studied. These other dimensions are important, because they may affect learning and school performance even in children with higher IQ. Older reports suggested associations between lead exposure and poorer classroom performance,7 impaired educational attainment,10,11 inattention and hyperactivity,12,13 juvenile delinquency,14 motor development,15–17 or behavioral problems.18–25 Lead was associated with externalizing (ie, aggression) and internalizing (ie, worry) problems in several studies using behavioral measures such as the Child Behavior Checklist.19,20,25

Studying behavior in the presence of an effect of lead on IQ presents both practical and inferential problems. First, understanding how lead exposure might affect the child’s psychological and emotional function at school and at home is approached both by questioning the parent and teacher about the child’s behavior and by testing specific relevant functional domains of mood and behavior. However, the tests may not isolate the specific domain of interest from IQ and in practice are often significantly correlated with IQ. Moreover, IQ may be on the pathway leading to behavioral problems.26 So, without appropriate control for IQ, the nature of the lead effect on behavior cannot be distinguished.25,27 Second, exposure to lead may continue from the fetal period through childhood. Although children’s lead concentrations correlate over time, the trajectories vary enough that a single measurement of blood concentration is insufficient to characterize an individual child’s exposure over time. In particular, we need to be able to distinguish whether peak blood lead level, which occurs at ~2 years old in the United States, or concurrent blood lead level, which is usually lower by school age when IQ and behavioral testing is done, accounts for any effect on behavior.

In the Treatment of Lead-Exposed Children (TLC) Study, we measured blood lead concentration periodically from 2 to 7 years old and have IQ, neuropsychological, and behavioral test scores at ages 5 and 7 years. Thus, we can analyze the association between lead and behavior while taking IQ into account and examine the relative strength of the association between blood lead concentrations at different ages, IQ, and behavioral test scores.

Subjects and Methods
The TLC study was a multicenter, randomized, placebo-controlled clinical trial of 780 children 12 to 33 months old (mean: 2 years; SD: 0.5 years), who had blood lead concentrations of 20 to 44 μg/dL, to investigate the effects of succimer, an oral chelating agent, on cognitive, behavioral, and physical development.28 The study was approved by the institutional review boards at the clinical centers, the Harvard School of Public Health, the Centers for Disease Control and Prevention, and the National Institute of Environmental Health Sciences. The parent(s) of all of the children provided written informed consent. Although ≤3 courses of treatment with succimer was effective in lowering blood lead concentrations for ∼9 to 10 months, it did not improve scores on tests of cognition, behavior, or neuropsychological function in children at 36 months29 or 60 months of follow-up30 as compared with placebo. Because the succimer treatment did not affect lead concentrations at baseline and ages 5 and 7 years, nor did it affect IQ and behavior scores at ages 5 and 7 years, the succimer and placebo study groups can be combined to study prospectively the effect of blood lead concentrations on the scores of neuropsychological and behavioral tests.

Blood Lead Concentrations
Venous blood was collected with lead-free containers twice before random assignment and on days 7, 28, and 42 after the beginning of each course of treatment. After treatment ended, blood lead concentrations were measured every 3 to 4 months. We use the second blood sample before random assignment (n = 780) as baseline (at age 2), the blood sample at 36 months of follow-up (n = 731) as the age 5 sample, and the last blood sample at 60 months of follow-up (n = 623) as the age 7 sample. The blood lead concentrations were measured at the Nutritional Biochemistry Branch of the Centers for Disease Control and Prevention by atomic absorption spectrometry based on the methods described by Miller et al.31 For blood lead concentrations at 7 years, 1 child who had a very high blood lead concentration of 51 μg/dL (~10 SDs from the mean) was excluded, leaving 622 for use in this analysis.

Cognitive Tests
At 5 years old, the child’s IQ was determined with the Wechsler Preschool and Primary Scales of Intelligence-Revised32; at 7 years old, child IQ was tested with the Wechsler Intelligence Scale for Children-III.33 At 1 of the visits between enrollment and the 36-month follow-up, the caregiver’s IQ (the mother for 88% of children, the father for 4%, and another caregiver for 8%) was measured with the 2 subtest versions of the Wechsler Adult Intelligence Scale-Revised.34,35
Behavioral Test Batteries

At 5 years of age, the Conners’ Parent Rating Scale-Revised (CPRS-R)\textsuperscript{36} was administered. The CPRS-R is a 27-item scale and yields an oppositional index, hyperactivity index, and attention-deficit/hyperactivity disorder (ADHD) index; the average of these 3 indices yields what we called the behavioral index.

At 7 years old, the children were tested with the Behavior Assessment System for Children (BASC) teacher rating scale (TRS) and BASC parent rating scale (PRS).\textsuperscript{37} The BASC for parents yields 4 composite scales: adaptive skills, behavioral symptoms index, externalizing problems, and internalizing problems. The BASC for teachers yields those 4 scales plus a school problems scale. Both CPRS-R and BASC yield T scores that have a mean of 50 and an SD of 10 in the general population. Higher CPRS-R and BASC scores generally indicated worse behavioral problems, except for the BASC adaptive skills scale, where higher scores were optimal.

Statistical Analysis

We examined the lead and behavioral associations while controlling for the lead effect on IQ. First we did a correlation analysis of behavior scores and concurrent IQ (ie, behavior at age 5 and IQ at age 5). Then we examined the lead and behavior associations using scatter plot and cubic smoothing splines (which provides a fitted curve constructed by piecewise polynomials) with S-PLUS software (Insightful Corp, Seattle, WA). Because spline regressions showed an approximately linear relation, we used linear models for examining the lead effect. Blood lead concentrations in TLC children were part of the eligibility criteria and, thus, have a restricted spread. This allows us to use the original (ie, untransformed) values of blood lead concentrations, facilitating the interpretation of the models. Based on the literature and our previous work with the data,\textsuperscript{19,24,25,38,39} a priori covariates include clinic center (Baltimore, MD; Newark, NJ; Philadelphia, PA; and Cincinnati, OH), race (black, white, and others), gender (male or female), language (English or Spanish), parent’s education (<12 years, 12 years, >12 years), parent’s employment (neither working or either working), single parent (yes or no), age at blood lead concentration test, and caregiver’s IQ. Treatment, per se, was not associated with behavior scores, and additional adjustment for treatment group did not markedly change the results and, thus, was not included in subsequent analyses.

We simultaneously estimated the strengths of lead effects on behavior, both the direct and indirect (through IQ; Fig 1), by using path analysis, a special case of structural equation modeling, which tests the fit of the correlation matrix against ≥2 causal models. In path analysis, a regression is done for each variable in the model (in our case, the behavioral test scores) as dependent on others that may be causal. When the model has ≥2 causal variables (in our case, lead and IQ plus covariates), path coefficients are partial regression coefficients that measure the extent of effect of 1 variable on another in the path model controlling for other previous variables. Path coefficients can be used to decompose associations into direct and indirect effects. A more detailed description of this procedure can be found in several references.\textsuperscript{40,41} The path analysis was tested via the LISREL 8 program\textsuperscript{42} using a maximum likelihood structural equation model, which provides unstandardized regression coefficients and their SEs for both direct and indirect effects.\textsuperscript{43} The total effect of lead on behavior would be the sum of the direct and indirect effects.

Blood lead concentrations in the same child are correlated, and multiple measures in 1 model are collinear and the coefficients are difficult to interpret. We, thus, constructed separate models for the different blood lead concentrations, either peak (at 2 years old) or concurrent with behavioral and IQ testing (5 or 7 years). All of the tests were 2-sided. Because of the difference in the number of children tested for each follow-up measurement, the sample sizes in the various regression models differ slightly.

For BASC scores at 7 years, we also did logistic regression on the percentage of children with BASC problem scores ≥60, including those at risk (score: 60–69) and with clinical behavioral problems (score ≥70), by concurrent blood lead concentration. Mplus software was used to calculate the direct and indirect effect of lead in the logistic models.\textsuperscript{44}

RESULTS

Four centers were involved in the recruitment, treatment, and follow-up of a total of 780 children in the TLC study: Baltimore (n = 213), Newark (n = 208), Philadelphia (n = 165), and Cincinnati (n = 194). A total of 396 children were randomly assigned to receive succimer and 384 to receive placebo. There were no differences between treatment and placebo groups in age, gender, race, and socioeconomic status (parents’ educa-
tion, employment, income, and receiving public assistance) at recruitment. Overall, the children were mostly black (77%), spoke English (95%), with a single parent (72%), and with parent(s) receiving public assistance (97%). Girls accounted for 44% of children, 40% of children had parents with <12 years of education, and 58% of children had neither parent employed.

Blood Lead Concentrations
At baseline, the mean blood lead concentration was 26 μg/dL. It declined to 12 μg/dL (range: 2–35 μg/dL) at 36 months of follow-up and to 8 μg/dL (range: 0–26 μg/dL) at 60 months of follow-up (Table 1). There were no differences in blood lead concentrations between succi-mer and placebo groups at these 3 age points (mean ± SD age: 2.0 ± 0.5, 5.0 ± 0.5, and 7.0 ± 0.2 years, respectively).

Cognitive Tests
The cognitive scores (mean ± SD) in TLC children at baseline, age 5, and age 7 are shown in Table 1. Care- giver’s IQ of these children had a mean of 80 and an SD of 11. Again, these cognitive scores of both children and caregivers did not differ by treatment group.

Behavioral Tests
The behavioral test scores (mean ± SD) at 5 and 7 years are shown in Table 1. Also shown are the correlation coefficients of these test scores with IQ and with blood lead concentration measured at the same age.

Lead and Behavior Association
We first plotted concurrent blood lead concentration and behavior scores and did spline regression. As examples, results of CPRS-R (at age 5) and teacher-rated BASC (at age 7) are shown in Figs 2 and 3. Behavioral problems tend to increase with increasing blood lead concentra- tion at both ages 5 and 7 in the unadjusted data, but the estimates of CPRS-R at age 5 and of BASC internalizing problems (both teacher and parent rated) at age 7 were not statistically significant. In the path analysis for the CPRS-R, there were no statistically significant direct effects of blood lead level at age 2 or age 5; indirect effects were small and not consistent (Table 2). At age 7, there were no statistically significant direct or indirect effects of blood lead concentration at age 2 (data not shown). Blood lead concentration at age 7 had a statistically significant direct effect on BASC-TRS behavioral symp-toms index, externalizing problems, and school prob- lems and BASC-PRS externalizing problems (Figs 4 and 5). There were indirect effects of blood lead concentra- tion at age 7 on all of the measurements except BASC-TRS externalizing problems and BASC-PRS internalizing problems. In the logistic regression analysis examining those at risk of or having clinically significant behavior problems, a 10 μg/dL elevation in 7-year lead was asso-

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Mean and SD of Blood Lead Concentrations, IQ, and Non-IQ Outcomes and the Correlation Coefficients Between IQ and Blood Lead Concentrations and Non-IQ Outcomes at Specific Age in TLC Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>Variables</td>
</tr>
<tr>
<td>2</td>
<td>Blood lead concentration, μg/dL</td>
</tr>
<tr>
<td>5</td>
<td>Blood lead concentration, μg/dL</td>
</tr>
<tr>
<td></td>
<td>IQ</td>
</tr>
<tr>
<td></td>
<td>CPRS-R</td>
</tr>
<tr>
<td></td>
<td>Oppositional index</td>
</tr>
<tr>
<td></td>
<td>Hyperactivity index</td>
</tr>
<tr>
<td></td>
<td>ADHD index</td>
</tr>
<tr>
<td></td>
<td>Behavioral index</td>
</tr>
<tr>
<td>7</td>
<td>Blood lead concentration, μg/dL</td>
</tr>
<tr>
<td></td>
<td>IQ</td>
</tr>
<tr>
<td></td>
<td>BASC-TRS</td>
</tr>
<tr>
<td></td>
<td>Adaptive skills</td>
</tr>
<tr>
<td></td>
<td>Behavioral symptoms index</td>
</tr>
<tr>
<td></td>
<td>Externalizing problems</td>
</tr>
<tr>
<td></td>
<td>Internalizing problems</td>
</tr>
<tr>
<td></td>
<td>School problems</td>
</tr>
<tr>
<td></td>
<td>BASC-PRS</td>
</tr>
<tr>
<td></td>
<td>Adaptive skills</td>
</tr>
<tr>
<td></td>
<td>Behavioral symptoms index</td>
</tr>
<tr>
<td></td>
<td>Externalizing problems</td>
</tr>
<tr>
<td></td>
<td>Internalizing problems</td>
</tr>
</tbody>
</table>

NA indicates not applicable.

<sup>a</sup> All P < .05 except for parent-rated internalizing problems at age 7.
<sup>b</sup> All P < .05 except for CPRS-R scores and BASC internalizing problems (both teacher and parent rated).
associated with increased risk in teacher-rated externalizing and school problems and parent-rated behavioral symptoms index by direct effect (Table 3).

DISCUSSION
In data from a clinical trial of lead-exposed children, we found that lead exposure was associated with behavior problems in urban 5- to 7-year-olds. Using a modeling strategy designed to separate direct and indirect effects of lead, we found that, in 5-year-olds, concurrent blood lead concentration had no direct effect on behavior, and indirect effects were small and inconsistent, although some were statistically significant. Thus, we believe that, if lead exposure is affecting behavior in 5-year-olds, it is doing so mostly though IQ, and direct effects have not emerged or are not measurable with the methods we used. Although the oppositional index in CPRS-R and parent-rated BASC externalizing problem scores were
correlated (with correlation coefficient = 0.49), these 2 indices may not capture the same behavioral dimension. In 7-year-olds, there is no effect of blood lead concentration at age 2. However, for the blood lead level at 7 years, there are direct, relatively large effects on the TRS of BASC (behavioral symptoms index, externalizing problems, and school problems); indirect effects on these are smaller (the indirect effect on externalizing problems was only borderline significant). For adaptive skills, the indirect effect is significant. Internalizing problems (excessive anxiety or worry) is the scale with the least effects, with a small but significant indirect effect and a similar but less precisely estimated direct effect. The results from the parent rating scale of the BASC are consistent with teacher report, showing a large direct effect on externalizing problems and a smaller but significant indirect effect on adaptive skills. In general, the results are consistent with a direct effect at age 7 of contemporaneously measured blood lead level on behavior, specifically, conduct and school problems, and an indirect effect through IQ on most other neuropsychological test scores.

The lead and IQ association has long been the focus of investigation for lead effects on the child’s nervous system, partly because of the easiness, reliability, and validity of IQ tests and the easy interpretation for both researchers and regulators. Noncognitive effects of lead, on the other hand, are much more complex to study. Furthermore, it has not been the norm in studies that did include behavioral or other measures to tease apart the lead effects on IQ to isolate a direct effect on behavior, even when IQ was measured. The Port Pirie cohort study reported that both externalizing and internalizing behavior problem scores were negatively associated with lead after controlling for child’s IQ: we find relatively large direct effects of lead on externalizing problems and smaller, indirect effects on internalizing problems.

In the studies of lead effects on child IQ, it has long been held that the cross-sectional association between lead and IQ in school-aged children could be the residual effects of peak blood lead concentration at approximately age 2 years. Recent analysis of TLC data and pooled analysis of 7 international cohort studies, how-

| TABLE 2 | Direct and Indirect Effects (95% Confidence Intervals) of 10 μg/dL Blood Lead Concentration on Behavioral Test Scores at Age 5 |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Behavioral Tests at Age 5 | Blood Lead at Age 2 | Direct | Indirect | Blood Lead at Age 5 | Direct | Indirect |
| CPRS-R | Oppositional index | 1.19 (−0.75 to 3.13) | 0.35 (0.02 to 0.68)* | 1.18 (−0.84 to 3.20) | 0.51 (−0.25 to 1.27) |
| Hyperactivity index | 0.93 (−0.89 to 2.75) | 0.34 (0.05 to 0.63)* | 1.10 (−0.80 to 3.00) | 0.50 (−0.83 to 1.83) |
| ADHD index | 0.54 (−1.17 to 2.25) | 0.61 (−0.39 to 1.61) | 0.54 (−1.22 to 2.30) | 0.90 (0.35 to 1.45)* |
| Behavioral index | 0.89 (−0.72 to 2.50) | 0.44 (−0.46 to 1.34) | 0.94 (−0.73 to 2.61) | 0.64 (0.23 to 1.05)* |

Data are adjusted for clinic center, race, gender, language, parent’s education, parent’s employment, single parent, exact age at blood lead concentration measurement, and caregiver’s IQ.

* P < .05.
ever, show that concurrent blood lead concentration has the strongest association with IQ scores. For behavior, such analyses are scarce; Burns et al.²⁵ found postnatal lead measures had associations with Child Behavior Checklist total behavior problem scores that were “qualitatively similar” to lifetime average lead exposure in the Port Pirie study. In our study, for both teacher and parent BASC scores, concurrent blood lead level generally had a stronger association than earlier blood lead measures. This is consistent with our previous analysis of the lead and IQ association.³⁹ The results suggest that prevention of lead exposure should continue into later childhood and not cease soon after peak blood lead level begins to fall at approximately age 3.

The biological mechanism of lead effects on cognitive function and neurobehavior has been studied for a long time. Lead has been found to affect synaptogenesis, postsynaptic N-methyl-D-aspartate receptor sensitivity, calcium-mediated events, neurotransmitter dopamine release, and mitochondria activities.⁴⁶–⁴⁸ However, the possible pathway of lead effects on behavior is still to be determined.

Our study has the strength of large sample size, long follow-up period, high retention rate of subjects in the follow-up, multiple measurements of behavior, and good quality control in the measurements. The limitations of the study are lack of Home Observation for Measurement of the Environment score, lack of some family and neighborhood characteristics (ie, social stressors), and limited generalizability to the general population because of the high blood lead level at enrollment (mean: 26 μg/dL in TLC study vs 2 μg/dL in US children).

**TABLE 3**

<table>
<thead>
<tr>
<th>Burt</th>
<th>BASC-TRS</th>
<th>BASC-PRS</th>
<th>BASC-TRS</th>
<th>BASC-PRS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% Score ≥60</td>
<td>OR (95% CI) per 10 μg/dL Concurrent Blood Lead</td>
<td>% Score ≥60</td>
<td>OR (95% CI) per 10 μg/dL Concurrent Blood Lead</td>
</tr>
<tr>
<td></td>
<td>Total Lead &lt;10 μg/dL</td>
<td>Lead ≥10 μg/dL</td>
<td>Direct Indirect</td>
<td>Total Lead &lt;10 μg/dL</td>
</tr>
<tr>
<td>BASC-TRS</td>
<td>Behavioral symptoms index</td>
<td>518</td>
<td>29</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Externalizing problems</td>
<td>517</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Internalizing problems</td>
<td>519</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>School problems</td>
<td>519</td>
<td>39</td>
<td>33</td>
</tr>
<tr>
<td>BASC-PRS</td>
<td>Behavioral symptoms index</td>
<td>620</td>
<td>33</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Externalizing problems</td>
<td>620</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Internalizing problems</td>
<td>620</td>
<td>17</td>
<td>17</td>
</tr>
</tbody>
</table>

Data were adjusted for clinic center, race, gender, language, parent’s education, parent’s employment, single parent, exact age at blood lead concentration measurement, and caregiver’s IQ.

* P < .05.
CONCLUSIONS
We have found, in children with relatively high lead exposure, that concurrent blood lead concentration was associated with externalizing and school problems at 7 years old, and the effect was not entirely mediated through the lead effect on IQ. On the other hand, higher blood lead concentration at 2 years of age was not associated with behavior at 7 years of age. Finding both direct and indirect effects of concurrent blood lead concentration on behavior among school-aged children lends further urgency to the necessity of preventing lead exposure in children, preferably continuing into school age.

ACKNOWLEDGMENTS
The TLC trial was supported by National Institute on Environmental Health Sciences intramural and extramural funds, in cooperation with the National Institutes of Health Office of Minority Health, and by the Centers for Disease Control and Prevention. Succimer and placebo capsules were gifts from McNeil Labs (Fort Washington, PA).

We thank David Dunson, PhD, at the National Institute of Environmental Health Sciences for helpful comments on an earlier version of the article.

REFERENCES
34. Wechsler D. *Wechsler Adult Intelligence Scale, Revised.* San Antonio, TX: The Psychological Corporation; 1981
39. 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
47. Dietrich KN, Ris MD, Succop PA, Berger OG, Bornschein RL. Early exposure to lead and juvenile delinquency. *Neurotoxicol Teratol.* 2001;23:511–518
Mortality of Late-Preterm (Near-Term) Newborns in Utah

Paul C. Young, MDa,b, Tiffany S. Glasgow, MDa,b, Xi Li, MStatc, Ginger Guest-Warnick, MSa,b, Gregory Stoddard, MPHc

Departments of Pediatrics and Internal Medicine and the Intermountain Consortium for Child Health Services Research, University of Utah School of Medicine, Salt Lake City, Utah

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

ABSTRACT

OBJECTIVES. The purpose of this work was to determine the relative risk for mortality and the causes and ages of death for late-preterm newborns (gestational age of 34–36 weeks) compared with those born at term.

METHODS. We reviewed data from birth and death certificates of infants born in Utah between 1999 and 2004. We calculated early neonatal (first week), neonatal (first 28 days), and infant (first year) mortality rates for each weekly estimated gestational age cohort from 34 to 42 weeks and, using 40 weeks as the reference, risk ratios for each cohort. Causes of death were grouped into 8 categories and compared for near term and term newborns. Crude mortality rates and risk ratios for death from all causes and for infants dying of all causes other than birth defects were measured.

RESULTS. Birth defects were the single-most common cause of death for both term and late-preterm newborns. Mortality rates for late-preterm newborns remained significantly higher after excluding those who died of birth defects from the comparisons.

CONCLUSIONS. Compared with those born at term, late-preterm (near-term) newborns have significantly higher mortality rates. Each weekly increase in estimated gestational age is associated with a decreasing risk of death. Birth defects are the leading cause of death among late-preterm newborns but do not entirely account for their higher risk of death.
Prematurity, defined as birth before 37 weeks’ gestation, is the major determinant of morbidity and mortality for newborns.1 Since the beginning of the modern era of newborn intensive care, the interest of neonatologists has focused primarily on attempting to improve the care of and outcomes for increasingly smaller and more premature newborns. In recent years, a subset of premature infants, those born between 34 and 37 weeks, has become the subject of increasing interest. In 2000, Kramer et al demonstrated that, in the United States and Canada, these infants contributed substantially to overall infant and neonatal mortality, although their mortality rate was significantly lower than that of newborns whose gestational age was <33 weeks. Because these infants represent ~75% of the total number of preterm infants, their deaths constitute a much larger “etiologic fraction” of infant and neonatal mortality than do those who are more premature.2 Others have pointed out that short-term morbidity, as reflected by increased hypoglycemia, jaundice, apnea, respiratory distress, longer lengths of stay, and higher costs, is also much greater for this cohort of infants.3 In addition, a recent report indicates that “moderately” low birth weight infants (1500–2500 g) are more likely, as they grow older, to develop a special health care need, a chronic condition, a learning disability, or attention-deficit disorder.4

The appropriate setting, NICU or term nursery, and the best approaches to providing care for these infants are subjects of increasing interest and controversy. To some extent this is reflected by the differing terminology used to describe those born before 37 weeks’ estimated gestational age (EGA). Some experts have suggested that the traditional designation “near-term” be replaced by “late preterm” to emphasize that it is preferable to approach these infants as “still preterm” rather “almost term.” The importance and increasing attention being paid to this group of infants is reflected by 2 national efforts. In July 2005, the National Institute of Child Health and Human Development invited a multidisciplinary team of experts to a workshop titled “Optimizing the Care and Outcome of the Near-Term Pregnancy and the Near-Term Newborn Infant.” In June 2005, the Association of Women’s Health, Obstetric and Neonatal Nurses launched a “multiyear initiative to address the unique physiologic and developmental needs of near-term infants.” Participants in both of these efforts emphasized the need for more empirical data to develop the appropriate evidence base to guide care. Although “near term” is probably in more common usage, in this article we will follow the recommendations of the National Institutes of Health consensus panel and refer to this group of newborns, those with gestational ages between 34 and 36 and 6/7 completed weeks, as “late preterm.”

Clinicians involved in the day-to-day care of late-preterm newborns, as well as those developing guidelines and recommendations, would benefit from having a clear understanding of the potential differences in risks faced by these infants compared with their more mature counterparts. We report the results of a study designed to quantify the differences in mortality and in the causes and patterns of death for late-preterm and term newborns. Knowledge of the degree of risk and information regarding the causes of death may be useful for developing recommendations for the initial assessment, monitoring, management, and follow-up for these newborns.

METHODS

To conduct our study, we analyzed data from birth and death certificates supplied by the Office of Vital Records and Statistics of the Utah Department of Health. We included all of the infants who were born in Utah between 1999 and 2004 with a gestational age ≥34 to ≤42 weeks. From death certificate data, we categorized the causes of death into 1 of the 8 groups proposed by the International Collaborative Effort on Birthweight, Perinatal and Infant Mortality (Table 1). For each weekly EGA cohort from 34 to 42 weeks, we calculated early neonatal (first week), neonatal (first 28 days), and infant (first year) mortality rates. Using 40 weeks as the reference, we used logistic regression to calculate odds ratios (ORs) for death for each weekly gestational age cohort for the 3 time periods and then generated risk ratios using the method proposed by Zhang and Yu.

Birth and death certificate data were obtained under Utah statute 26-2-22 and Administrative Rule R436-17, which allow the release of birth and death data for qualified research studies. The study was determined to be exempt by the institutional review board of the University of Utah Health Sciences Center.

RESULTS

There were 283,975 births of infants with an EGA ≥34 to ≤42 weeks in Utah between 1999 and 2004. Of these, 651 died in the first year. Because the exact age at death for 2 of these newborns could not be determined, they were excluded from the analyses of early neonatal and neonatal mortality. Tables 2, 3, and 4 provide the crude mortality rates and risk ratios with 95% confidence intervals (CIs) for dying in the early neonatal, neonatal, and infant time periods, respectively, for infants from 34 to 42 weeks. Compared with those born at 40 weeks, mortality rates were significantly higher for all 3 of the periods for newborns who were born at 34, 35, 36, and 37 weeks. Neonatal and infant mortality rates were also slightly higher for those born at 38 weeks, whereas early neonatal and infant mortality rates were slightly higher for those born at 42 weeks.

As shown in Table 5, birth defects were the leading cause of death for newborns of all gestational ages for all 3 of the mortality periods. The impact of birth defects as a cause of death was greatest in the late-preterm infants.
Of the 166 late-preterm newborns who died during the first year, the cause of death in 105 (63%) was a birth defect, whereas for those with an EGA ≥37 weeks, birth defects accounted for 43% of the infant deaths. Analyses performed after removing those infants who died of birth defects continued to show higher early neonatal, neonatal, and infant mortality rates for the late-preterm infants (Tables 6–8). Table 9 displays the individual causes of death for each of the gestational age cohorts and ORs for dying of the particular cause relative to dying of that cause at 40 weeks’ EGA. Taking into account that some of the cells on the table contain small numbers, in general, newborns born at 34 to 36 weeks were more likely than those born at 40 weeks to die in the first year from all causes except infections and external causes. It was not possible to obtain valid comparisons of the individual causes of death by EGA cohort for...
the early neonatal and neonatal periods because of the small numbers in many of the cells.

We performed analyses for 2 other time periods, “late neonatal” (days 8–28) and “postneonatal” (days 29–365). Mortality rates for infants born at 34, 35, and 36 weeks were significantly greater for each cohort as a whole and after excluding those who died of birth defects (data not shown).

DISCUSSION

We found that mortality rates for newborns who were born after gestations of 34, 35, and 36 weeks were significantly greater compared with those born at 40 weeks. Even after removing from the analyses those who died of birth defects, the leading cause of death for these infants, the higher mortality rates remained. Compared with those born at 40 weeks, these infants were more likely to die of 6 of the 8 international collaborative effort mortality categories. The lack of an association between deaths from infection and external causes should be interpreted with caution, because there were few deaths from these causes in our data set. We also found a slight but statistically significant increase in mortality rates for some of the time periods for those born at 37, 38, and 42 weeks compared with those born at 40 weeks. The differences in relative risk for mortality for newborns of these 3 gestational ages compared with those born at 40 weeks are small, but, because of the large numbers of infants born at these gestational ages, particularly if extrapolated to include the entire US population, they may contribute substantially to mortality rates. We believe that this finding may be important to consider when planning the timing of delivery of otherwise uncomplicated pregnancies.

Our findings are consistent with those of Kramer et al1 and offer further detail by characterizing mortality rates for each week of gestational age. The differences between those born at 34, 35, and 36 weeks suggest that it is inappropriate to regard late-preterm infants as a homogeneous group with respect to mortality.

Our data are unable to address one of the most important questions: why the mothers of these infants delivered them before term. We do not know how many of these mothers were induced before term because of the presence of a complicating condition, such as pregnancy-induced hypertension, infection, or diabetes. We also do not know if any of these mothers received tocolytic agents in an effort to delay delivery. Although the appropriate end point for tocolysis is a subject of controversy among obstetricians,10–12 we suspect that, for some of these mothers, efforts to delay delivery were discontinued once the gestational age of their infant reached 34 weeks. The association of birth defects with both late-preterm delivery and mortality suggests that, in some cases at least, the presence of a lethal birth defect may have resulted in preterm labor and delivery. However, the disadvantage associated with delivery at 34, 35, or 36 weeks remained significant even when those infants who died of birth defects were excluded from the analysis.

At the investigators’ institution, the University of Utah Hospital, newborns with an EGA ≥34 weeks who appear stable are typically admitted to the well-baby nursery, where members of the Division of General Pediatrics manage their care. Newborns with an EGA of 34 completed weeks or those of any EGA who require intensive care are admitted to the NICU. The investigators had noted that the management of newborns with an EGA of 37 weeks in our well-baby nursery was highly variable. As clinicians responsible for directing the care of these infants, we recognize that, regardless of gestational age, the care of every infant must be individualized. However, we also believe that evidence-based guidelines for the initial assessment and management, discharge criteria, and the short-term and long-term follow-up of these late-preterm newborns may enable us to recognize potential risks that are modifiable and to treat them. Our data suggest that, even in the absence of a potentially lethal birth defect, late-preterm infants demand more careful evaluation and follow-up throughout their first year. The data of Stein et al4 although based on birth weights of 1500 to 2500 g rather than gestational age and, therefore, not strictly comparable, suggest that, because special health care needs and chronic conditions are more common, careful long-term follow-up is also appropriate.

The increased mortality rates for late-preterm newborns have substantial and increasing public health implications. According to the Institute of Medicine (IOM), the rate of premature births in the United States has risen >30% since the 1980s, and a significant majority (~75%) of premature births are in the gestational ages defined as near-term or late preterm.13 The reasons for the increase in premature births are unknown but probably include increasing maternal age; increased success of fertility treatments, which may result in multiple
births; and environmental, genetic, and other factors. The IOM report calls for a substantial increase in research to determine the causes of prematurity and to improve birth outcomes.

There are several important limitations to our data. Gestational ages, causes, and ages of death were based on birth and death certificate data. The method of obtaining gestational age estimates varies from hospital to hospital but is usually based on a combination of interviewing the mother and reviewing the medical chart; it is likely that there are some inaccuracies. Utah data may differ from that of other areas in the country. Although Utah’s population includes a substantial and increasing Latino component, most newborns are white; there are few black residents. We did not analyze our data by race or ethnicity, but other studies would suggest that prematurity rates and adverse outcomes for nonwhite populations are higher, suggesting that our findings may actually underestimate the risks associated with late-preterm birth. Finally, we do not know the causes of the preterm deliveries. It is likely that the causes of early delivery differ between those late-preterm infants who died and those who did not; such information will be very important for future studies assessing specific risk factors for late-preterm newborns.

Despite the limitations, we believe that our findings have importance for clinicians and researchers. Although the majority of late-preterm or near-term newborns do well, clinicians need to recognize that, as a

<table>
<thead>
<tr>
<th>Gestational Age, wk</th>
<th>Mortality Ratesa</th>
<th>Risk Ratios</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>1.27</td>
<td>3.25b</td>
<td>1.4–7.7</td>
</tr>
<tr>
<td>35</td>
<td>1.06</td>
<td>3.21b</td>
<td>1.5–6.1</td>
</tr>
<tr>
<td>36</td>
<td>0.66</td>
<td>2.69b</td>
<td>1.5–4.7</td>
</tr>
<tr>
<td>37</td>
<td>0.08</td>
<td>1.91b</td>
<td>1.2–3.1</td>
</tr>
<tr>
<td>38</td>
<td>0.15</td>
<td>1.95b</td>
<td>1.3–2.9</td>
</tr>
<tr>
<td>39</td>
<td>0.08</td>
<td>1.02</td>
<td>0.7–1.6</td>
</tr>
<tr>
<td>40</td>
<td>0.08</td>
<td>Ref</td>
<td>—</td>
</tr>
<tr>
<td>41</td>
<td>0.07</td>
<td>1.32</td>
<td>0.7–1.5</td>
</tr>
<tr>
<td>42</td>
<td>0.00</td>
<td>4.76b</td>
<td>1.5–15.4</td>
</tr>
</tbody>
</table>

a Mortality rates are per 1000 live births.
b Ratios are significantly more than 40 weeks.

<table>
<thead>
<tr>
<th>Gestational Age, wk</th>
<th>Mortality Ratesa</th>
<th>Risk Ratios</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>3.79</td>
<td>4.62b</td>
<td>2.5–8.4</td>
</tr>
<tr>
<td>35</td>
<td>3.16</td>
<td>3.91b</td>
<td>2.3–6.5</td>
</tr>
<tr>
<td>36</td>
<td>2.55</td>
<td>3.14b</td>
<td>2.0–4.9</td>
</tr>
<tr>
<td>37</td>
<td>1.45</td>
<td>1.81b</td>
<td>1.2–2.8</td>
</tr>
<tr>
<td>38</td>
<td>1.41</td>
<td>1.75b</td>
<td>1.2–2.5</td>
</tr>
<tr>
<td>39</td>
<td>0.82</td>
<td>1.02</td>
<td>0.7–1.5</td>
</tr>
<tr>
<td>40</td>
<td>0.80</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>41</td>
<td>0.90</td>
<td>1.13</td>
<td>0.6–2.1</td>
</tr>
<tr>
<td>42</td>
<td>2.79</td>
<td>3.42b</td>
<td>1.1–10.4</td>
</tr>
</tbody>
</table>

a Mortality rates are per 1000 live births.
b Ratios are significantly more than 40 weeks.

dating gestational age estimates varies from hospital to hospital but is usually based on a combination of interviewing the mother and reviewing the medical chart; it is likely that there are some inaccuracies. Utah data may differ from that of other areas in the country. Although Utah’s population includes a substantial and increasing Latino component, most newborns are white; there are few black residents. We did not analyze our data by race or ethnicity, but other studies would suggest that prematurity rates and adverse outcomes for nonwhite populations are higher, suggesting that our findings may actually underestimate the risks associated with late-preterm birth. Finally, we do not know the causes of the preterm deliveries. It is likely that the causes of early delivery differ between those late-preterm infants who died and those who did not; such information will be very important for future studies assessing specific risk factors for late-preterm newborns.

Despite the limitations, we believe that our findings have importance for clinicians and researchers. Although the majority of late-preterm or near-term newborns do well, clinicians need to recognize that, as a
7.6 (2.4–24.3)
40
2.1 (0.45–9.7)
50
3.5 (1.3–9.1)
77
1.4 (0.50–3.8)
88
1.9 (0.91–4.2) 149
0.73 (0.31–1.7) 143
Ref
72
0.84 (0.18–3.8) 25
5.6 (0.71–43.7)
7
—
651

ACKNOWLEDGMENT
We thankfully acknowledge the assistance of Jeff Duncan (director, Office of Vital Records and Statistics, Utah
Department of Health).

REFERENCES
For causes, see Table 1. SIDS indicates sudden infant death syndrome; Ref, reference; NA, no deaths in this category (cannot calculate ORs).
a ORs are shown with 95% CIs in parentheses relative to 40 weeks.
b Proportion of deaths from this cause differs from 40 weeks.

ORa
N (%)

4 (10.0)b
2 (10.0)
7 (9.1)b
6 (6.8)
20 (13.4)
12 (8.4)
10 (13.9)
2 (8.0)
1 (14.3)
64 (9.8)
NA
NA
0.62 (0.08–5.0)
1.44 (0.47–4.4)
0.85 (0.31–2.4)
0.60 (0.23–1.6)
Ref
1.05 (0.22–4.9)
7.0 (0.88–55.9)
—

ORa
N (%)
ORa
N (%)
ORa
N (%)
ORa
N (%)
ORa
N (%)
ORa
N (%)
ORa
N (%)

YOUNG et al

34 28 (70)b 22.2 (12.9–38.3) 2 (5.0)b 19.0 (2.7–135.1) 4 (10.0)b 12.7 (3.6–45.0) 0 (0.0)
NA
1 (2.5) 6.3 (0.66–60.1) 1 (2.5)
2 (0.3–19) 0 (0.0)
35 32 (64)b 13.7 (8.0–23.3) 7 (14.0)b 37.0 (7.7–178.3) 1 (2.0)
1.8 (0.21–14.6) 0 (0.0)
NA
3 (6.0)b 10.6 (2.1–52.4) 5 (10.0)b 7 (2–20)
0 (0.0)
5 (6.5)b 4.1 (1.26–13.5) 1 (1.3) 0.45 (0.06–3.5) 5 (6.5)b 8.3 (1.97–34.5) 7 (9.1)b 4 (2–11)
1 (1.3)
36 46 (60)b 9.5 (5.8–15.6) 5 (6.5)b 12.4 (2.4–63.8)
37 50 (57)b 4.8 (2.9–7.8)
3 (3.4)
3.4 (0.58–20.6) 2 (2.3)
0.77 (0.15–3.8) 4 (4.5) 0.84 (0.27–2.6) 8 (9.1)b 6.1 (1.6–23.1) 10 (11.4)b 3 (1.1–7) 5 (5.7)
38 62 (42)b 2.5 (1.6–4.0)
4 (2.7)
1.95 (0.36–10.6) 9 (6.0)
1.46 (0.52–4.1) 7 (4.7) 0.62 (0.24–1.6) 11 (7.4) 3.6 (1.0–12.8) 29 (19.5)b 4 (1.6–8) 7 (4.7)
39 62 (43) 1.5 (0.96–2.5) 3 (2.1)
0.91 (0.15–5.4) 11 (7.7)
1.11 (0.41–3.0) 10 (7.0) 0.55 (0.23–1.3) 7 (4.9) 1.4 (0.36–5.5) 30 (21.0)b 22 (1.1–5) 8 (5.6)
40 24 (33)
Ref
2 (2.8)
Ref
6 (8.3)
Ref
11 (15.3)
Ref
3 (4.2)
Ref
8 (11.1)
Ref
8 (11.1)
41 12 (48)b 2.1 (1.05–4.2) 1 (4.0)
2.1 (0.19–23.1) 0 (0.0)
NA
3 (12.0) 1.14 (0.32–4.1) 1 (4.0) 1.4 (0.15–13.4) 4 (16.0)b 2 (1.1–5) 2 (8.0)
42
4 (57)b 9.3 (3.2–26.9) 0 (0.0)
NA
0 (0.0)
NA
1 (14.3) 5.1 (0.66–39.4) 0 (0.0)
NA
0 (0.0)
NA
1 (14.3)
Total 320 (49)
—
27 (4.1)
—
38 (5.8)
—
37 (5.7)
—
39 (6.0)
—
94 (14.4)
—
32 (4.9)

Total
Remaining
External
SIDS
Other Speciﬁc Causes
Infections
Asphyxia
Immaturity
Birth Defects

EGA,
wk

TABLE 9 Causes of Death for Each Gestational Age Cohort in the First Year
e664

group, they are at increased risk of dying throughout the
first year. Although our study did not address short-term
or long-term morbidity, other studies suggest that both
are increased.3,4 Pediatricians and others who are responsible for both the initial and long-term care of latepreterm newborns must be aware of these increased
risks. Our findings provide clear evidence of a substantial increased risk of mortality for this group of infants
who have, to this point, been relatively ignored by the
research community. The March of Dimes, the IOM,
the National Institute of Child Health and Human Development, and Association of Women’s Health, Obstetric and Neonatal Nurses have all called for more
research regarding the causes and appropriate management of premature newborns, including those defined
as near term or late preterm.6,7,13,15 We believe that our
study provides important evidence to support these initiatives.

The contribution of mild and moderate preterm birth to infant
mortality. Fetal and Infant Health Study Group of the Canadian
age distribution among U.S. singleton births: impact on rates of
3. Wang ML, Dorer DJ, Fleming MP, Catlin EA. Clinical outcomes
4. Stein RE, Siegel MJ, Bauman LJ. Are children of moderately
low birth weight at increased risk for poor health? A new look
5. Engle WA. A recommendation for the definition of “late preterm” (near-term) and the birth weight-gestational age classification system. Semin Perinatol. 2006;30:2–7
6. Raju TN. Introduction. Semin Perinatol. 2006;30:1
collaborative effort (ICE) on birth weight, plurality, perinatal,
and infant mortality. III: a method of grouping underlying
causes of infant death to aid international comparisons. Acta
JAMA. 1998;280:1690 –1691
10. Jones SC, Brost BC, Brehm WT. Should intravenous tocolysis
be considered beyond 34 weeks’ gestation? Am J Obstet Gynecol.
2000;183:356 –360
11. Lewis DF, Futayyeh S, Towers CV, Asrat T, Edwards MS,
Brooks GG. Preterm delivery from 34 to 37 weeks of gestation:
1996;174:525–528




Transmission of West Nile Virus Through Human Breast Milk Seems to Be Rare

Alison F. Hinckley, PhD, Daniel R. O’Leary, DVM, Edward B. Hayes, MD

Arboviral Diseases Branch, Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, US Public Health Service, Department of Health and Human Services, Fort Collins, Colorado

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

ABSTRACT

INTRODUCTION. In September 2002, possible transmission of West Nile virus via human milk was reported for the first time.

METHODS. Since 2003, the Centers for Disease Control and Prevention collected reports of maternal or infant West Nile virus illness during the breastfeeding period. All of the reported instances were reviewed. In addition, milk samples from women infected during pregnancy were tested for West Nile virus RNA and West Nile virus–specific antibodies.

RESULTS. Six infants were reported to have breastfed from mothers with West Nile virus fever. Five of the 6 infants had no illness or detectable antibodies to West Nile virus in serum after onset of maternal illness. One infant who was not tested and developed a rash was otherwise well 1 week after onset of maternal illness. In addition, 2 infants were reported to have developed West Nile virus illness while breastfeeding; preceding maternal illness was not documented. Two breastfed infants whose mothers acquired West Nile virus fever in the last week of pregnancy developed West Nile virus–specific antibodies; both infant infections could have been congenitally acquired. Of 45 milk samples from women infected with West Nile virus during pregnancy, 2 had West Nile virus RNA, and 14 had immunoglobulin M antibodies to West Nile virus.

CONCLUSIONS. Of 10 reported instances since 2003 of maternal or infant West Nile virus illness while breastfeeding, transmission of West Nile virus through human milk could neither be ruled out nor confirmed for 5 cases; in 5 others, serologic tests indicated no vertical transmission. Transmission of West Nile virus through breastfeeding seems to be rare, but more information is needed.
WEST NILE VIRUS (WNV) has become a major public health threat in the United States, causing ≥8000 cases of human neuroinvasive disease since its detection in New York in 1999. WNV is a flavivirus transmitted to humans by infected mosquitoes. In 2002, it was discovered that WNV could be transmitted through breast transfusion, organ transplantation, and transplacentally. In September 2002, the first case of possible WNV transmission via human milk was reported. Because the majority (~80%) of persons infected by WNV are asymptomatic, cases where transmission occurs via breastfeeding may go unrecognized.

Other arthropod-borne flaviviruses can be transmitted to humans and animals through ingestion of milk or milk products. In humans, infection with and illness because of tick-borne encephalitis virus has occurred after consumption of raw sheep, goat, and cow’s milk and sour-milk products made of infected milk. In animals, louping-ill and Powassan viruses have been transmitted from lactating goats to suckling kids. In 1 study, transmission of louping-ill virus through breast milk resulted in infection and clinical illness of 5 of 13 milk-exposed goat kids.

The first possible transmission of WNV via human milk was reported to the Centers for Disease Control and Prevention (CDC) in September 2002. In this case, a 40-year-old woman was transfused immediately postpartum with blood that was subsequently found to contain WNV nucleic acid. She began to breastfeed on the day of delivery and continued through day 16 postdelivery, her second day of hospitalization for WNV encephalitis. A sample of undiluted mother’s milk from day 16 postdelivery tested positive for the presence of WNV nucleic acid and WNV-specific immunoglobulin M (IgM) and IgG antibodies. Viral culture of the milk was negative for WNV. A second sample of undiluted milk collected 24 days after delivery was negative for WNV nucleic acid. This sample was positive for WNV-specific IgM antibodies when tested at a 1:400 dilution. At 25 days of age, serum from the breastfed infant tested WNV-specific IgM positive, although the child remained healthy. The mother reported that the infant had been kept primarily indoors, with no obvious exposures to mosquitoes. Below we summarize available surveillance data regarding other cases of possible WNV infection via breastfeeding.

METHODS
Since 2003, data on WNV disease in humans, including demographic and clinical information on women or children who became ill while breastfeeding, have been routinely reported to the CDC by 54 state and local health departments through ArboNET, a national arboviral disease surveillance system. Also in 2003, CDC and state health departments organized a surveillance registry to assess pregnancy and infant health outcomes for women infected with WNV during pregnancy. This enhanced surveillance system was designed to collect maternal and infant health information and diagnostic samples, including human colostrum or mature human milk samples. Breast milk and serum specimens collected from case subjects were tested at a 1:400 dilution for WNV-specific IgA (breast milk only), IgM, and IgG antibodies by an enzyme-linked immunosorbent assay (ELISA). Positive ELISA results were confirmed by plaque-reduction neutralization tests for neutralizing antibodies to WNV, St Louis Encephalitis virus (SLEV), and other closely related flaviviruses. To indicate specificity to WNV, only samples with neutralizing antibody titers ≥10 and 4 times higher than corresponding titers to St Louis Encephalitis virus were considered positive. Whenever possible, undiluted cerebrospinal fluid (CSF) was tested using these same techniques. Breast milk and CSF were also tested for the presence of WNV nucleic acid by polymerase chain reaction (PCR). These activities were defined by the National Center for Infectious Diseases as public health practice, thus institutional review board regulations do not apply, and informed consent was not required.

RESULTS
Mothers Who Were Breastfeeding at Time of Maternal WNV Illness
In August of 2003, 4 mothers who developed WNV fever while breastfeeding were reported to the CDC. All 4 of the mothers had IgM to WNV detected in serum shortly after illness onset. In the first instance, a breastfeeding infant developed a rash and no other symptoms 11 days after onset of WNV illness in the mother. No infant serum or maternal breast milk samples were collected; therefore, the possibility of transmission of WNV by breastfeeding could not be further evaluated. In the second instance, a breast milk sample obtained 12 days after the onset of maternal WNV illness tested positive for WNV-specific IgA antibodies; negative for WNV-specific IgM, IgG, and neutralizing antibodies; and negative for the presence of WNV RNA. The 4-month-old infant remained healthy and had no detectable antibodies to WNV in serum collected 13 days after the onset of maternal illness. In the third instance, a breast milk sample collected 46 days after the onset of maternal WNV illness tested negative for IgA, IgM, IgG, and neutralizing antibodies and negative for the presence of WNV RNA. The 2-month-old infant remained healthy and had no detectable WNV antibodies in serum collected nearly 7 weeks after maternal illness. In the last instance, a breast milk sample obtained ~5 months (192 days) after the onset of maternal WNV illness tested negative for WNV-specific IgA, IgM, IgG, and neutralizing antibodies and negative for the presence of WNV RNA. The newborn infant remained healthy and had no detectable WNV RNA.
specific antibodies in serum collected 11 days after the mother’s breast milk sample was collected.

No similar reports were received during 2004, but in 2005, 2 more mothers who developed WNV fever while breastfeeding were reported to the CDC. Again, both mothers developed IgM to WNV shortly after illness onset. Breast milk samples were obtained from the first mother at 44, 45, 47, and 48 days after the onset of WNV illness. All of these samples tested negative for WNV-specific IgA, IgM, and neutralizing antibodies, as well as for WNV RNA. This woman’s 7-month-old infant remained healthy and had no detectable antibodies to WNV in serum collected at 48 days after the onset of maternal illness. Breast milk samples collected from the second mother at 28, 29, 31, 32, and 35 days after the onset of WNV illness were also negative for WNV RNA but were not tested for WNV-specific antibodies. This woman’s 12-month-old infant remained healthy and had no detectable antibodies to WNV in serum collected at 36 days after the onset of maternal illness. Information pertaining to these mothers who were breastfeeding at time of maternal WNV illness is summarized in Table 1.

Infants Who Were Breastfeeding at Time of Infant WNV Illness

In 2003, the CDC received reports of 2 infants who were breastfeeding at the time of their WNV illness. The first was a girl who developed fever and rash at 8 months of age. Serum obtained 1 week after the onset was negative for WNV-specific IgM antibodies. The mother did not report illness of life, coarctation of the aorta, and a bicuspid aortic valve. Cord blood, infant serum, andcolostrum were not

and maternal serum was subsequently positive for IgM antibodies to WNV. Because the incubation period for symptomatic WNV infections is thought to be shorter than 15 days,15 and this mother’s illness followed the onset of her child’s illness by nearly a month, it seems unlikely that transmission through breastfeeding occurred unless the mother was asymptotically infected with WNV while the child was breastfeeding, and her subsequent symptomatic illness was of different etiology.

The second breastfed infant developed a fever and rash that progressed to encephalitis at 9 months of age. Serum and CSF specimens were positive for WNV-specific IgM antibodies. The mother did not report illness before the infant’s illness; asymptomatic WNV infection of the mother could not be evaluated, because no maternal specimens were collected.

Breastfeeding Infants With Mothers Infected Before Delivery

Two previously reported women who had been enrolled in the CDC’s WNV pregnancy registry in 2003 gave birth to infants who breastfed and subsequently tested positive for WNV IgM antibodies, indicating WNV infection of the infants.11 However, maternal infection occurred close to parturition, making it difficult to know whether transmission occurred before or after delivery. The first mother became acutely ill with WNV fever and rash 5 days before delivery in August 2003. Her serum was positive for WNV-specific IgM antibodies 7 days after her illness onset. Her male child was born afebrile with a transient maculopapular rash that resolved at 36 hours of life, coarctation of the aorta, and a bicuspid aortic valve. Cord blood, infant serum, andcolostrum were not
obtained at birth. Infant serum collected 2 months after birth was positive for WNV-specific IgM and neutralizing antibodies. A second infant serum sample collected at ~8 months of age was positive for IgM and IgG to WNV, and the neutralizing antibody titer to WNV was fourfold higher than the titer in the previous sample. The child remained otherwise healthy. The mother’s breast milk sample obtained 20 days after delivery was negative for the presence of WNV RNA and for IgA, IgM, and IgG antibodies to WNV but was positive for neutralizing antibodies. Another breast milk sample collected 2 months after delivery was negative for all of the WNV-specific antibodies. Because of the late identification of infant WNV infection relative to birth, neither congenital transmission nor transmission by mosquito exposure can be excluded.

The second mother had onset of WNV illness and rash in September 2003, 6 days before delivery. A maternal serum sample collected 12 days after illness onset tested positive for WNV-specific IgM, IgG, and neutralizing antibodies and was negative for WNV RNA. A cord blood sample collected at delivery, 6 days after the onset of symptoms, tested negative for WNV-specific IgM, IgG, and neutralizing antibodies. At 10 days of age, the infant girl developed neonatal meningitis, and her CSF and serum were both positive for WNV-specific IgM antibodies. Breast milk obtained 37 days after onset of the mother’s illness tested positive for WNV-specific IgM antibodies (testing for IgA was not performed) and neutralizing antibodies and negative for WNV RNA.

**WNV RNA and Antibodies in Breast Milk**

Thirty-two colostrum and 13 mature human milk samples collected from 45 of 72 women enrolled in the CDC’s WNV pregnancy registry during the 2003 and 2004 transmission seasons were tested for WNV RNA by PCR. Two colostrum samples were positive; all of the other samples were negative. One positive sample was collected 50 days after the onset of maternal illness; the other sample was collected 70 days after the onset. Because of the low level of RNA detected in these samples, virus isolation was not attempted. The 2 children exposed to these colostrum samples did not apparently become infected with WNV, as evidenced by the lack of WNV-specific antibodies in blood samples drawn at 291 and 308 days of age, respectively. For both cases where colostrum was positive, WNV RNA was not detected in maternal serum at ~2 months after the onset of illness. Of the 43 WNV PCR-negative specimens, 9 were collected <70 days after the onset. The number of milk samples testing positive for the presence of RNA or WNV-specific antibodies is presented in Table 2.

Thirty-four colostrum and 13 breast milk samples were tested for WNV-specific IgM antibodies. Eleven colostrum samples (32%) and 3 breast milk samples (23%), collected at 8, 23, and 31 days postpartum, respectively, tested positive. The average time from onset of illness to collection of IgM-positive breast milk and colostrum samples was 81 days (range: 24–153 days). Time from onset of illness to collection of IgM-negative samples was longer, with an average of 184 days (range: 11–293 days; P < .0001, by unpaired t test).

Six (19%) of 31 colostrum samples and 1 (11%) of 9 breast milk samples tested for WNV-specific IgA antibodies were positive. The average time from onset of illness to collection of IgA-positive breast and colostrum samples was 92 days (range: 11–158 days). Time from onset of illness to collection of IgA-negative samples was longer, with an average of 159 days (range: 26–293 days; P = .04).

All 31 colostrum and 11 breast milk samples tested for IgG to WNV by ELISA were negative. However, neutralizing antibodies, which may be IgM, IgG, or IgA, against WNV were found in 26 (84%) of 31 colostrum samples and 6 (55%) of 11 breast milk samples. The average time from onset of illness to collection of samples testing positive for neutralizing antibodies was 118 days (range: 11–281 days) compared with 170 days (range: 37–293 days) for negative samples (P = .07).

**DISCUSSION**

Only 1 likely instance of WNV transmission by breastfeeding has been reported to date, in 2002, when the mother was infected postdelivery by transfusion with blood products containing WNV. Through heightened surveillance, we found no evidence of WNV transmission to infants of breastfeeding mothers with laboratory-documented postpartum WNV illness and were unable to assess possible transmission in a sixth infant. Two other reported infants were breastfeeding at the time of their WNV infection, but preceding maternal WNV infection was not documented in either case. Lastly, 2

<table>
<thead>
<tr>
<th>Specimen</th>
<th>RNA Positive, n (%)</th>
<th>Total, N</th>
<th>IgM Positive, n (%)</th>
<th>Total, N</th>
<th>IgA Positive, n (%)</th>
<th>Total, N</th>
<th>IgG Positive, n (%)</th>
<th>Total, N</th>
<th>Neutralizing Antibodies Positive, n (%)</th>
<th>Total, N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast milk</td>
<td>0 (0)</td>
<td>13</td>
<td>3 (23)</td>
<td>13</td>
<td>1 (11)</td>
<td>9</td>
<td>0 (0)</td>
<td>11</td>
<td>6 (55)</td>
<td>11</td>
</tr>
<tr>
<td>Colostrum</td>
<td>2 (6)</td>
<td>32</td>
<td>11 (33)</td>
<td>34</td>
<td>6 (19)</td>
<td>31</td>
<td>0 (0)</td>
<td>31</td>
<td>26 (84)</td>
<td>31</td>
</tr>
<tr>
<td>All</td>
<td>2 (4)</td>
<td>45</td>
<td>14 (30)</td>
<td>47</td>
<td>7 (18)</td>
<td>40</td>
<td>0 (0)</td>
<td>42</td>
<td>32 (76)</td>
<td>42</td>
</tr>
</tbody>
</table>
other infants became infected while breastfeeding after maternal infection that occurred before delivery. These infants could have acquired their infection congenitally, through breastfeeding, or from infected mosquitoes.

Several differences exist between the mother thought to have transmitted WNV to her infant through breast milk in 2002 and the breastfeeding mothers in subsequent years who were ill but apparently did not transmit the virus. The mother in 2002 acquired infection through blood transfusion and developed WNV encephalitis, whereas the other mothers acquired infection from infected mosquitoes and developed WNV fever. Either a higher virus inoculum from blood transfusion or host factors in the mother at the time of infection, including asymptomatic mastitis or unreported minor areolar lesions, might have increased the likelihood of breast milk transmission.

Immune factors in human breast milk and those in the infant gastrointestinal tract and saliva may interact synergistically to provide protective antiviral benefits. Antiviral activity against 2 flaviviruses, Japanese B encephalitis virus and Dengue virus, have been observed in the fat fraction of human milk. Milk lipids provide antiviral activity during digestion, when they are broken down into fatty acids and monoglycerides by both milk-derived bile-salt–stimulated lipase and lipolytic activity in the infant’s gut.

In addition, secretory IgA, the predominant immunoglobulin found in human milk, is thought to protect the recipient infant from common enteric and respiratory pathogens. In addition to the antiviral agents in human milk that were already mentioned, several other components, including lactoferrin, lysozyme, glycoconjugates, and oligosaccharides, may contribute to protection against flavivirus infection.

Infant age may also influence the likelihood of acquiring WNV infection from human breast milk. For example, the intestinal immaturity of the neonate can allow for increased attachment of certain pathogenic organisms to gut epithelial cells. Epithelial barriers in the newborn’s intestinal tract are stimulated by growth factors in human milk, including epidermal growth factor, lactoferrin, cortisol, and hormones. Later in postnatal development, IgA-producing plasma cells appear in the infant’s intestinal tract. An infant’s systemic immune response is also age dependent, with serum concentrations of IgM rising gradually after 6 days of age, and maternally acquired IgG declining gradually during the first 6 to 8 months of life.

Although no human colostrum and mature human milk samples tested in this study at a 1:400 dilution were positive by the ELISA test for WNV-specific IgG, 76% of all of the milk samples tested neutralized WNV in the plaque-reduction neutralization test assay, reflecting relatively low ELISA test sensitivity with breast milk specimens, the presence of non-IgG virus-neutralizing antibody, and/or the presence of other nonimmunoglobulin virus-neutralizing components in the milk. Proportionally more colostrum samples than mature human milk samples contained WNV-specific IgM and IgA antibodies, which may reflect the shorter time from onset of illness to sample collection for these first postpartum milk samples and is consistent with expected immunoglobulin content in human milk in the first days postpartum. According to our data, WNV-specific IgM antibodies were detected more frequently in human milk after WNV infections than IgA antibodies. This may be because of the absence or transient nature of an IgA immune response to WNV infection or limitations in the assay’s ability to detect WNV-specific IgA.

Because flaviviral RNA would be expected to decrease in tissues with increasing time after illness, WNV RNA may be more likely found in colostrum than breast milk from a woman who was infected with WNV during pregnancy. Furthermore, there are higher concentrations of macrophages in colostrum as compared with mature human milk, and WNV is thought to target monocytes and macrophages for the replication and spread of infection throughout tissues. Our data suggest that the presence of WNV RNA in milk seems to be rare after maternal WNV infection during pregnancy (only 2 colostrum samples were positive of 45 total milk samples). For both cases where colostrum was positive, WNV RNA was not detected in serum at ~2 months after the onset of illness. Although no information has been published regarding the persistence of flavivirus nucleic acid in human breast tissues, tick-borne encephalitis virus has been isolated from brain tissues of animals and humans at >7 months after infection, and Powassan virus has been isolated from the brain of a fatal human case at 42 days after the onset of illness, indicating that arthropod-borne flaviviruses may sometimes persist in human tissues.

In this investigation, only 6 women and 4 infants who were breastfeeding at the time of their WNV infection were reported to the CDC. It is likely that these numbers are far below the actual number of breastfeeding women and infants who were infected with WNV since the 2002 transmission season. Thus, the information obtained from this sample of cases may not be representative of the entire pool of breastfeeding women or infants infected with WNV. Another limitation is the lack of knowledge regarding the validity of laboratory serology techniques used to detect WNV-specific antibodies in human milk, because these tests have not been standardized for milk samples. Variability in collection, shipment, and storage of these human milk specimens may have also compromised the ability to accurately measure immunoglobulins and RNA using standard testing techniques.

This article reports results for those few women or infants having clinical WNV illness who were reported to a national surveillance system by health care providers.
and health departments. The results do not support any change in infant feeding practices after infection with WNV. Additional information is needed to understand the potential for WNV transmission through breast milk. For areas where WNV is endemic, clinicians and public health authorities should include WNV in the differential diagnosis of illness among breastfeeding children and nursing mothers. The CDC continues to seek information and offer testing of specimens, including milk and blood samples, from women diagnosed with WNV while nursing.

ACKNOWLEDGMENTS
We give special thanks to Stephanie Kuhn for assistance with case identification and specimen procurement. We also thank the state WNV surveillance coordinators and the following CDC staff: Theresa Smith and Roy Campbell for general advice; Robert Lanciotti, Roselyn Hochbein, Amanda Noga, Amy Lambert, Olga Kosoy, Janeen Laven, and Brandy Russell for laboratory analysis of specimens; and Peggy Collins for database development/management.

REFERENCES
27. Weingartl HM, Neufeld JL, Copps J, Marszel P. Experimental West Nile virus infection in blue jays (Cyanocitta cristata) and crows (Corvus brachyrhynchos). Vet Pathol. 2004;41:362–370
Influence of Multiple Risk Behaviors on Physical Activity–Related Injuries in Adolescents

Ian Janssen, PhD, Suzanne Dostaler, MSc, William F. Boyce, PhD, William Pickett, PhD

Department of Community Health and Epidemiology, School of Kinesiology and Health Studies, Department of Emergency Medicine, Social Program Evaluation Group, Faculty of Education, and Centre for Health Services Policy and Research, Queen’s University, Kingston, Ontario, Canada

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

ABSTRACT

OBJECTIVE. The study objectives were to (1) examine the relationship between physical activity and physical activity injuries in youth, (2) determine whether this relationship is modified by the engagement in multiple risk behaviors, and (3) determine whether this relationship is modified by the setting of the injury (school versus outside of school).

METHODS. We examined associations between physical activity and multiple risk behaviors with physical activity injuries occurring at and outside of school. The study population consisted of a representative sample of 5559 Canadian youth in grades 6 through 10 who participated in the 2001/2002 Health Behavior in School-Aged Children Survey. The exposure and outcome measures were determined from a classroom-based survey.

RESULTS. Irrespective of grade, there were strong gradient relations between physical activity participation and related injuries outside of school. Conversely, there were modest relations between physical activity participation and related injuries at school. In students in grades 6 to 8, there was no relation between multiple risk behaviors and physical activity injuries at school and a curvilinear relation between multiple risk behaviors and physical activity injuries outside of school. The opposite pattern of relationships between multiple risk behaviors and injuries was observed in students in grades 9 to 10. Irrespective of grade and setting of injury, there was no significant interaction between physical activity and multiple risk behaviors on injury risk. The results were consistent by severity of injury and for structured/organized and unstructured/informal forms of physical activity.

CONCLUSIONS. The environment moderated the relation between physical activity and related injuries in that strong risk gradients only existed outside of the school setting. Unexpectedly, there were no consistent gradients between the engagement in multiple risk behaviors and physical activity injuries or any interaction effect between physical activity exposure and multiple risk behaviors. These findings suggest that optimizing the environment would be a preferred strategy for preventing physical activity injuries compared with selectively targeting youth who engage in multiple risk behaviors.
Physical activity is associated with numerous health benefits in young people including the maintenance of a healthy body weight, an improved cardiovascular disease risk profile, reduced reports of depression, and a high level of confidence and self-esteem. Given the impact of physical activity on these outcomes and the increasingly sedentary life lead by many young people, increasing the physical activity level in children and adolescents has become the focus of public health initiatives across North America and Europe.

One of the only negative impacts of physical activity participation is the increased risk of injuries that occurs while participating in sports. For instance, in a representative sample of 11- to 15-year-old Canadians, 64% of the medically treated injuries occurred while participants were engaged in sports and other forms of physical activity. Injuries are a leading cause of morbidity and the leading cause of mortality in young people. Thus, injury prevention is also a leading public health concern for pediatric populations.

An important issue to address is how to increase sport and physical activity participation while simultaneously minimizing injury risk. Some obvious strategies are to increase the use and effectiveness of protective equipment and to modify the rules of a sport in a way that would reduce injury occurrence. An alternative approach to these strategies is to identify youth who are at particularly high risk of having a physical activity injury and who should be targeted for behavior modification (eg, change their risk behaviors).

Along that line of thought, a recent theme of research in adolescent injury etiology has examined the relation between the engagement in multiple risk behaviors (MRBs) and injury. MRB refers to the engagement in activities such as smoking and drug use, sensation seeking (eg, doing dangerous things just for fun), and failure to take appropriate safety precautions (eg, not wearing a helmet while riding a bicycle). It has consistently been reported that youth who engage in MRBs are at increased injury risk compared with youth who do not engage in these behaviors. Although we are unaware of any studies that have targeted sports injuries, it is reasonable to assume that young people who engage in risky behaviors are more aggressive risk takers in sport and, subsequently, are at greater risk of having a physical activity–related injury.

There is also interest in the features of the environment that influence injury risk in children and adolescents. School environments, for example, are influential settings for the establishment of social norms, such as risk-taking behaviors. Furthermore, optimization of the physical environment has the potential to minimize injury risk. Thus, one may expect to see a lower physical activity injury risk in a more controlled environment, such as the school.

The primary study objectives were to (1) examine the relation between physical activity and the risk of physical activity injuries in young people, (2) determine whether this relationship is modified by the engagement in MRBs, and (3) determine whether this relationship is modified by the setting of the injury (school versus outside of school). We hypothesized that the risk of injury for a given amount of physical activity would be higher in youth who engaged in MRBs compared with youth who do not engage in these behaviors. We also hypothesized that the relationship among physical activity, MRBs, and injuries would be less pronounced in a controlled environment, such as the school.

METHODS

Description of Survey and Study Population
Results are based on the Canadian records from the 2001/2002 World Health Organization Health Behavior in School-Aged Children Survey (HBSC). The HBSC is a cross-sectional survey from 35 countries. The survey consisted of a classroom-based questionnaire. The Canadian data were collected in the first half of 2002. The sample was designed according to the international HBSC protocol in that a cluster design was used with the school class being the basic cluster, the distribution of the students reflected the distribution of Canadians in grades 6 to 10, and the sample was self-weighted. Samples were selected to represent distributions of schools by size, location, language, and religion. A total of 74.2% of the students selected for the study completed the questionnaire. The total sample consisted of 7266 children. We excluded those who did not respond to the exposures and outcomes of interest, which left a total of 5559.

The Canadian HBSC was approved by the Queen’s University General Research Ethics Board. Consent was obtained from the participating school boards, individual schools, parents, and students. Student participation was voluntary.

Exposure Variables

Physical Activity Participation
Two physical activity exposures were considered: the amount of physical activity performed at school and outside of school. Participants were asked 4 questions regarding the number of hours per week that they usually participated in moderate-to-vigorous-intensity physical activity (eg, activities that make you out of breath or warm) in class time at school, in their free time at school, outside of school while participating in lessons or league or team sports, and outside of school while participating in informal activities either alone or with friends. For each of these questions there were 9 response categories ranging from “none at all” through “7 hours or more.” The number of hours for questions 1 and 2 were summed to determine the amount of physical activity performed at school. The num-
number of hours for questions 3 and 4 were summed to determine the amount of physical activity performed outside of school. In addition to being treated as continuous variables, the physical activity scores were placed into low (<3 hours/week), moderate (3–6 hours/week), and high (≥7 hours/week) categories. These categories were chosen to represent <30 minutes, 30 to 60 minutes, and ≥60 minutes of physical activity per average day. Furthermore, the 2 at-school (class time and free time) and 2 outside-of-school (organized sport and informal activities) physical activity measures were considered as separate exposures to determine whether the results were consistent by type of physical activity in each setting.

**MRBs**

Common risk behaviors were documented as follows: current smoking (4 categories: “smoke every day” through “I do not smoke”), lifetime drunkenness (5 categories: “never” through “≥10 times”), current use of seatbelts (5 categories: “always” through “never”), lifetime cannabis use (7 categories: “never” through “40 or more times”), illicit drug use (eg, ecstasy, cocaine, or lysergic acid diethylamide) during lifetime (7 categories: “never” through “40 or more times”), nonuse of condoms during most recent sexual intercourse (3 categories: “I have never had sexual intercourse,” “yes,” or “no”). The questions on cannabis use, illicit drug use, and condom use were only asked to grades 9 to 10 students. Responses for each item were given a point score corresponding with frequency of engagement (“never” = 0, “occasional” = 1, “frequent” = 2). The points were summed for the various items to create an MRB score (range: 0–6 points for grades 6–8 and 0–12 points for grades 9–10). In addition to being treated as a continuous variable, the MRB score was grouped into “never” (0 points in grades 6–8 and grades 9–10), “occasional” (1–3 points in grades 6–8 and 1–6 points in grades 9–10), and “frequent” (4–6 points in grades 6–8 and 7–12 points in grades 9–10) categories. The MRB scale used here was validated previously in this cohort using a confirmatory factor analysis.29

**Outcome Variables**

The 2 outcomes examined were self-report of a medically treated (by doctor or nurse) injury that occurred while participating in physical activity at school in the 12 months before the survey and self-report of a medically treated injury that occurred while participating in physical activity away from school in the previous 12 months. An initial question determined whether the student was injured in the past year to a degree that required medical treatment. Using a series of questions, individuals reporting ≥1 medically treated injury were asked about activities leading to their most serious injury and where they were when this injury happened. Individuals who reported that the most serious injury occurred while playing or training for a sports/recreational activity, biking/cycling, riding a skate scooter, skating (roller blades, skateboard, or ice skating), or walking/running were considered to have a physical activity injury. Injuries that occurred at school, including the school grounds, were considered to be school injuries. Injuries that occurred at home/in the yard, at a sports facility or field, in the street/road/parking lot, in the countryside (such as a lake, park, etc), or other locations (eg, commercial or business area) were considered to be nonschool injuries. To consider whether the results differed by injury severity, analyses were repeated after creating subcategories of severe (met any of the following: missed ≥1 full day of school, had an operation, or stayed in hospital overnight) and nonsevere injuries (all other injuries).30

**Statistical Analysis**

All of the analyses were conducted using SAS (SAS Institute, Cary, NC). The initial set of analyses considered the bivariate relation between the primary exposures (physical activity at school and outside of school, MRBs) and outcomes (physical activity injuries at school and outside of school) using logistic regression models. The second set of analyses considered the multivariate relation between the exposures and outcomes using logistic regression models. A product term (physical activity × MRB) was also included in the multivariate models to determine whether there was an interaction effect for the 2 exposures. Physical activity and MRB scores were included in the regression models as continuous variables. Full cubic polynomial models were used to determine whether the relations were linear or nonlinear in nature. Gender was treated as a covariate in all of the logistic regression models after analyses determined that gender was not an effect modifier in the relation between the exposures and outcomes. The constant and β coefficient parameters of the logistic regression models were used to calculate the probability of having a physical activity injury, which is the outcome presented in the results.

**RESULTS**

**Descriptive Information**

Table 1 describes the study sample. Within both the grade 6 to 8 and 9 to 10 groups, ~8% suffered a medically treated injury while participating in physical activity at school and ~30% suffered a medically treated injury while participating in physical activity outside of school in the past year. The study participants averaged ~5.5 hours of physical activity per week both at school and outside of school. For grade 6 to 8 youth, 24.6% and 24.8% were in the “low at school” and “low outside of school” physical activity categories, 36.3% and 35.0% were in the “moderate at school” and “moderate outside of school” physical activity categories, and 39.1% and
40.1% were in the “high at school” and “high outside of school” physical activity categories. The corresponding values for grade 9 to 10 students were 28.2%, 27.4%, 36.6%, 34.1%, 35.2%, and 38.5%, respectively. Regarding the MRB score, the majority of youth were considered to be never or occasional risk takers, with only 4.6% of grade 6 to 8 and 13.3% of grade 9 to 10 students being considered frequent risk takers (Table 2). More extensive details on the engagement in individual risk behaviors are contained within Table 2.

Relation Between Physical Activity and Physical Activity Injuries

The relation between physical activity level and the probability of having a physical activity injury that required medical attention within the past year is illustrated in Fig 1. For physical activity injuries that occurred at school, there were modest positive relations for students in grades 6 to 8 ($P_{\text{linear}} = .07$) and grades 9 to 10 ($P_{\text{linear}} = .04$). For physical activity injuries that occurred outside of school, there were strong gradient (or dose–response) effects for youth in grades 6 to 8 ($P_{\text{linear}} < .001$) and grades 9 to 10 ($P_{\text{quadratic}} = .04$). In addition to using physical activity as a continuous variable, these analyses were performed using physical activity groups (Table 3). Regardless of grade, the likelihood of having a physical activity injury outside of school was increased in the moderate and high physical activity groups by comparison with the low physical activity groups. Significant relations were not observed for physical activity injuries occurring at school.

The above analyses were repeated after physical activity injuries were subclassified into severe and nonsevere injuries. Without exception, the results were consistent by injury severity. The above analyses were also repeated after subclassifying school physical activity into class time and free time and physical activity outside of school into organized sport and informal activities. Again, the results were consistent with those reported in the preceding paragraph.

Relation Between MRBs and Physical Activity Injuries

Figure 2 illustrates the probability of having a medically treated physical activity injury within the past year according to MRB score. In grade 6 to 8 students, there was no significant ($P = .25$) relation between MRBs and physical activity injuries occurring at school. There was a curvilinear relation ($P_{\text{quadratic}} = .01$) between MRBs and physical activity injuries occurring outside of school such that the probability of injury was lowest in those who engaged in no risky behaviors, highest in those who occasionally engaged in some risky behaviors, and low in those who frequently engaged in all risky behaviors. In grade 9 to 10 students, the relation between the MRB score and physical activity injuries occurring at school was curvilinear in nature ($P_{\text{quadratic}} = .02$) such that the probability of injury was low in those who engaged in no risky behaviors, highest in those who occasionally engaged in risky behaviors, and lowest in those who frequently engaged in all risky behaviors. There was no significant ($P = .64$) relation between the MRB score and physical activity–related injuries occurring out-

### Table 1: Descriptive Characteristics of Study Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Grades 6–8 (n = 3846)</th>
<th>Grades 9–10 (n = 1713)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>54.0</td>
<td>57.8</td>
</tr>
<tr>
<td>Male</td>
<td>46.0</td>
<td>42.2</td>
</tr>
<tr>
<td>Severe injuries while participating in physical activity, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class time at school</td>
<td>1.7</td>
<td>3.3</td>
</tr>
<tr>
<td>Free time at school</td>
<td>2.6</td>
<td>1.7</td>
</tr>
<tr>
<td>Organized sport outside of school</td>
<td>6.7</td>
<td>11.1</td>
</tr>
<tr>
<td>Informal activities outside of school</td>
<td>7.1</td>
<td>6.1</td>
</tr>
</tbody>
</table>

### Table 2: Prevalence (%) of Engagement in Risky Behaviors

<table>
<thead>
<tr>
<th>Risky Behavior</th>
<th>Grades 6–8 (n = 3846)</th>
<th>Grades 9–10 (n = 1713)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>Occasional</td>
<td>Frequent</td>
</tr>
<tr>
<td>Multiple risk behavior scale</td>
<td>62.6</td>
<td>32.8</td>
</tr>
<tr>
<td>Smoking</td>
<td>92.9</td>
<td>3.5</td>
</tr>
<tr>
<td>Drunkenness</td>
<td>84.0</td>
<td>8.4</td>
</tr>
<tr>
<td>Nonuse of seatbelts</td>
<td>72.0</td>
<td>17.8</td>
</tr>
<tr>
<td>Cannabis use</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Other drug use</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Nonuse of condoms</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Data are presented as prevalences. — indicates that the risky behavior was not assessed in this age group.
side of school in grade 9 to 10 youth. These analyses shown in Fig 2 were also performed using MRB groups (Table 4). The likelihood of having a physical activity injury was not significantly higher in the occasional and frequent MRB groups, with the exception of physical activity injuries occurring outside of school for grade 6 to 8 students.

The results in the proceeding paragraph were consistent by injury severity subcategories with the exception of the following. In grade 6 to 8 students, there was a positive relation between MRB score and severe physical activity injuries occurring outside of school, whereas the relation between MRB score and nonsevere injuries was nonsignificant. The above analyses were also consistent by subcategory of school (class time versus free time) and outside of school (organized sport versus informal activities) activities with 1 exception. In grade 9 and 10 students, the relation between MRBs and physical activity injuries occurring outside of school was positive for informal activities and negative for organized sport.

Interactive Effect of Physical Activity and MRBs on Physical Activity Injuries

The next set of analyses considered the interactive effect of physical activity level and MRBs on physical activity injuries. Irrespective of grade and location of injury, physical activity and MRB levels predicted injury; however, there was no significant interaction between physical activity and MRBs on the injury outcomes. An example of this effect is illustrated in Fig 3, which shows the gradient relation between physical activity level and physical activity injuries occurring outside of school according to level of engagement in MRBs. The intercepts of the regression lines between physical activity and injury varied by MRB level, but the slopes of the regression lines did not. This indicates that, for a given level of physical activity, the injury risk varied according to MRB level; however, the modifying effects of MRB did not vary across the physical activity range. For instance, in grade 9 to 10 students, the absolute difference in the probability of a physical activity injury occurring outside of school between never and occasional MRB participants was 9.3% for those who spent an average of 1 hour of per week participating in physical activity, 11.6% for those who spent an average of 7 hours per week participating in physical activity, and 12.1% for those who spent an average of 14 hours per week participating in physical activity. These differences were modest and not statistically significant.

A second example illustrating the lack of interaction effect on physical activity injuries is illustrated in Fig 4. This figure depicts the relation between the MRB score and physical activity injuries that occurred at school in low, moderate, and high physical activity groups. Regardless of physical activity level, the relation between
MRBs and physical activity injuries occurring at school was not significant \((P > .2)\).

**DISCUSSION**

This article presents a novel analysis of the correlates of physical activity injuries in adolescents. The findings indicate that gradients exist between the amount of physical activity that adolescents engage in and the occurrence of physical activity injuries. The social environment (at school versus outside of school) moderated the relation between physical activity and related injuries in that strong risk gradients only existed outside of the school setting. Unexpectedly, there were no consistent gradients between the engagement in MRBs and physical activity injuries or any interaction effect between physical activity exposure and MRB. Thus, there was little support for the hypothesis that youth who engage in risky lifestyle behaviors are at particularly increased likelihood of having a self-reported physical activity injury after controlling for the exposure to activity.

The population health framework suggests that risks for health outcomes, such as injuries, are driven by interactions between individual and contextual risk factors.\(^{31}\) To our knowledge, this is the first study to examine individual risk factors while simultaneously examining the roles of settings (contextual risks) as correlates of physical activity injuries in young people. Interestingly, there was strong evidence that moderate injury risk gradients in the relations between physical activity participation and related injuries were much weaker in schools settings. This observation was consistent for severe and nonsevere injuries and for structured/organized and unstructured/informal forms of physical activity. We had no information on the quality of the school environments responsible for this observation, although we speculate that this is, in part, explained by contextual factors and, in particular, enforced rules and expectations around student behavior and the supervision provided by adult role models that is present in the school setting. The Centers for Disease Control and Prevention have published a set of evidence-based school guidelines for injury prevention that speak to these variables.\(^{27}\) Some of the key recommendations in the Centers for Disease Control and Prevention guidelines are to establish social and physical environments that promote safety, implement health and safety education curricula and instruction, provide safe physical education and extracurricular physical activity programs, and provide staff development services that impart knowledge to promote safety.

An earlier study led us to believe that strong gradients would exist between the engagement in MRBs and physical activity injuries in adolescents. In an analysis of the 1997/1998 version of the Canadian HBSC, Pickett et al\(^{23}\) found a strong gradient relation between the number of risk-taking behaviors (of a possible 7 risky behav-
iors) that a young person engaged in and his or her likelihood of having a sports injury. That study did not control for the amount of physical activity that the participants engaged in, which is a limitation given that young athletes, at least at the college level, tend to partake in more risky lifestyle behaviors than their peers. Furthermore, that study used a crude additive measure of MRBs as compared with a confirmed and factor analytically derived MRB scale that was used here. Finally, the analytical approaches were different in the 2 studies in that Pickett et al used logistic regression that focused on categories of risk-taking behavior as exposures, whereas in the present study, we used a continuous risk behavior scale that smoothes out a mathematical function.

When we did see relations between MRBs and physical activity injuries in this study, they tended to be weak and inconsistent by grade. For instance, MRBs were not related to physical activity injuries occurring outside of the school setting in grade 9 to 10 students, whereas for grade 6 to 8 students, there was a curvilinear relation between the MRB score and physical activity injuries occurring outside of school. These age differences may be related to the different variables used to construct the risk behavior scales in grade 6 to 8 versus grade 9 to 10 students, the greater participation in most risky behav-

**FIGURE 3**
Probability of having a medically treated physical activity–related injury outside of school in the past year according to the average amount of physical activity performed outside of school in the average week. Data are grouped into never-, occasional-, and high-MRB categories (see “Methods”). The data lines represent the best-fit regression lines calculated from the logistic regression models.

**FIGURE 4**
Probability of having a medically treated physical activity–related injury at school in the past year according to the MRB score. Data are grouped into low, moderate, and high physical activity categories (see “Methods”). The data lines represent the best-fit regression lines calculated from the logistic regression models.
iors in the older adolescents, and/or different types of physical activity that younger and older adolescents may engage in. On the basis of the data collected, it is not possible to identify the reasons underlying the observed relationships, and these interpretations, therefore, remain speculative.

It is noteworthy that we used a global measure of risk-taking behavior that was assessed from questions on substance abuse (tobacco, alcohol, and drugs) and safety precautions (seat belt use and condom use). Thus, we did not assess risk-taking behaviors in sports and physical activity, per se, because they were not available within the survey. These risks would include things such as sliding head first in baseball or diving to head a ball in soccer. In an earlier study that considered a 9-item soccer risk-taking behavior scale, Kontos33 found that risk-taking behaviors on the soccer field were not directly related to soccer injuries in a sample of 260 soccer players aged 11 to 14 years. Consistent with these findings, a recent systematic review of risk-taking behaviors and injuries concluded that sports-related risk-taking behaviors are not associated with sports injuries in adults.34 It has been postulated that athletes who partake in thrill seeking and other risks and who are confident in their abilities are better able to handle the risks, are more likely to engage in calculated risks, and are subsequently less likely to get injured.33,34 Injuries that occur while participating in sports seem to be the only type of injuries that are not related to risk-taking behavior.34

Our results suggest that there is an increased risk of injury among youth engaging in physical activity outside of the school setting. The more time spent in activities, the greater the injury risk. These results are consistent with a wide body of literature12–14 and indicate that the risks associated with physical activity need to be considered when planning sports and other activities. Contrary to our original hypothesis, there is little evidence here that targeting high-risk youth (as inferred by their engagement in risk-taking behaviors) for behavioral modification as a primary prevention strategy would have a meaningful impact on decreasing injury rates as opposed to more general strategies aimed at the minimization of risks in physical activities and associated environments. With that being said, it is important to note that the present study was not an evaluation of behavior modification to reduce injury in high-risk youth, and there is a well-recognized risk of injury during physical activity among youth who engage in MRBs. In fact, all youth who engaged in physical activity, regardless of their engagement in risky behaviors, may need more guidance in injury prevention, particularly for activity performed outside of the school setting.

There are a number of strengths of this study including the size and scope of the analyses, the use of standard measures and survey procedures, and the anonymous nature of reporting, which should promote accuracy in responses. Limitations of this study include the use of self-reported data and the cross-sectional nature of the survey, which limits the ability to infer causal relationships about the findings. Second, although the HBSC questionnaire items have undergone extensive piloting and validation,28 the possibility of biased reporting of risky behaviors motivated by a desire to provide socially desirable responses must be recognized, as well as the potential for error in the recall of injuries over the past year. Third, the 5559 completed surveys used in our analyses represent only 57% of the potential youth who could have responded, which is further indicative of a potential responder bias introduced because of the high level of nonresponse. Finally, the HBSC survey did not collect information on whether or not the students were wearing appropriate safety equipment when they were injured, the specific times at which the physical activity injuries occurred, and the specific sports/activities the students were participating in when the physical activity injuries occurred. Information on these variables may have added further clarification to the relationships examined in this study.

CONCLUSIONS

The purpose of this study was to examine relationships among physical activity, adolescent risk-taking, and the occurrence of physical activity injuries at school and outside of school. Contrary to expectations, there was little evidence that youth who engage in MRBs should be selectively targeted as a primary prevention strategy for physical activity injuries. However, there was strong evidence that schools are protective settings. That is, whereas engagement in physical activity was associated with physical activity–related injuries in a graded, dose-related manner outside of school, this was not the case for the more controlled school setting. This suggests there is something about the school setting that acts as a mediator of the risk behavior-injury relationship. Although not characterized here, this mediator is most likely to involve enforced rules and expectations around student behavior and the supervision provided by adult role models that is present in the school setting. Injury prevention specialists must look to other prevention strategies to minimize risks in nonschool settings. These could include optimization of the physical environment, establishment of norms surrounding peer-group behavior, and more comprehensive strategies of adult supervision.

ACKNOWLEDGMENTS

This study was supported by research agreements with the Canadian Institutes of Health Research (operating grant 2004MEP-CHI-128223-C) and the Public Health Agency of Canada (contract HT089-05205/001/SS), which funds the Canadian version of the World Health Organization-HBSC. The World Health Organization-
REFERENCES


HBSC is a World Health Organization/Euro collaborative study. The international coordinator of the 2001/2002 study was Candace Currie (University of Edinburgh, Edinburgh, Scotland), and the databank manager was Oddrun Samdal (University of Bergen, Bergen, Norway).
Neurodevelopmental Functioning in HIV-Infected Infants and Young Children Before and After the Introduction of Protease Inhibitor–Based Highly Active Antiretroviral Therapy

Jane C. Lindsey, ScD®, Kathleen M. Malee, PhD®, Pim Brouwers, PhD®, Michael D. Hughes, PhD®, for the PACTG 219C Study Team

ABSTRACT

OBJECTIVES. The purpose of this work was to examine the effects of HIV infection and the impact of highly active antiretroviral treatment with protease inhibitors on neurodevelopmental functioning during the first 3 years of life.

PATIENTS AND METHODS. Pediatric AIDS Clinical Trials Group 219/219C is a longitudinal cohort study that has enrolled HIV-infected (HIV+/H11001) and HIV-exposed but uninfected (HIV+/H11002) infants and children since 1993. Longitudinal profiles of neurodevelopmental functioning as measured by the Bayley Scales of Infant Development were compared by HIV-infection status before and after the availability of highly active antiretroviral therapy with a protease inhibitor and within infants with Bayley tests available before and after initiating protease inhibitor therapy.

RESULTS. In the pre–protease inhibitor era, mean mental and motor scores in HIV+ (n = 54) infants <1 year of age were significantly lower than those among HIV− infants (n = 221) and remained lower up to 2 years of age. After protease inhibitors became available, mean mental and motor functioning of HIV+ infants (n = 91) <1 year of age were still significantly lower than those of HIV− infants (n = 838). However, against a background of declining scores among the HIV− infants, there was evidence of limited improvement in the HIV+ infants relative to their uninfected peers. Among infants who had Bayley II evaluations before and after starting a protease inhibitor, there was a trend to improved mental and motor scores after initiation of protease inhibitor therapy.

CONCLUSIONS. The suppression of systemic viral replication and subsequent substantial improvements in survival and immunologic status brought about by highly active antiretroviral therapy have been followed by limited improvements in neurodevelopmental functioning in young children. Additional longitudinal research is needed to better understand the role of antiretroviral therapy as well as the impact of genetic and environmental factors on neurodevelopmental functioning in children affected by HIV.
Infection with HIV is associated with an increased risk for central nervous system (CNS) disease, primarily because of HIV infection in the brain. Early in the epidemic, 50% to 90% of children with HIV infection exhibited severe and often progressive CNS manifestations, reflected in deficits in cognitive, language, motor, and behavioral functioning. In the absence of treatment, the greatest risk for HIV-associated encephalopathy occurs during the first year of life and may be the initial AIDS-defining symptom, but CNS manifestations of HIV may also first present years later.

Recent investigations have suggested a reduced prevalence of encephalopathy in HIV-infected (HIV+) children. However the differential impact of factors responsible for this decline remains unclear. Use of highly active antiretroviral therapy (HAART) in children has dramatically improved survival and growth. Suppression of systemic viral replication may reduce the number of HIV-infected cells entering the CNS and potentially reduce the incidence of severe CNS damage. However, many antiretroviral agents, including some protease inhibitors (PIs), are not equivalent in their ability to penetrate the blood brain barrier, which may allow the CNS to serve as a reservoir for latently infectious virus. Since HAART became available, a proportional increase in the AIDS dementia complex compared with other AIDS-defining illnesses has been observed in HIV+ adults. Therefore, HAART may have less impact on HIV-related CNS dysfunction and neurodevelopmental functioning. A recent small case series described significant cognitive decline among children treated with HAART in the context of immunologic, virologic, and clinical stability.

The purpose of this investigation was to examine the effects of HIV infection on neurodevelopmental functioning during the first 3 years of life and to evaluate the impact of PI-containing HAART regimens. We report on neurodevelopmental functioning of HIV-exposed but uninfected (HIV−) and HIV+ infants and young children, regardless of treatment regimen, before and after PIs became available. We also examine the within-subject effects of initiation of a PI before 3 years of age.

METHODS

Study Design

These analyses used data from Pediatric AIDS Clinical Trials Group (PACTG) Late Outcomes Protocol 219/219C. This prospective cohort study, which opened to accrual in 1993, was designed to assess long-term outcomes of in utero exposure to HIV and antiretroviral drugs in infants and children, as well as the long-term effects of antiretroviral therapy (ART) among children with HIV infection. Until 2000, only children enrolled in PACTG clinical trials or children born to HIV-infected mothers enrolled in PACTG or Adult AIDS Clinical Trials Group trials were eligible. In 2000, eligibility criteria for PACTG 219 were expanded (PACTG 219C) to include all perinatally exposed infants and HIV+ children, regardless of previous participation in a clinical trial. Human subject research review boards at participating sites in the United States, including Puerto Rico, approved the research protocol. Before enrollment, written informed consent was obtained from participants’ parents or legal guardians.

Developmental Measures

The Bayley Scales of Infant Development (B-I) and the Bayley Scales of Infant Development, Second Edition, were used to assess the developmental functioning of infants. The B-I was used in PACTG 219 predominately in the pre-PI era, before 1997, and can be administered to infants and children 1 month and 24 days of age to 30 months and 15 days. PACTG 219 transitioned to the B-II starting in March of 1996, and its use continued in PACTG 219C. The B-II can be administered to infants and children 16 days of age to 42 months and 15 days. As sites transitioned to using the B-II, they were encouraged to continue using the B-I test in infants who had already completed a B-I test to avoid test transition effects. Tests were administered at entry to the study and then every 6 months until 2 years of age and, for the B-II, at 3 years. Only tests completed by infants within the age limits of the test were included in our analyses.

Results from both Bayley tests were recorded as raw scores and then transformed using the normative tables in the manuals to age-adjusted scaled scores with no adjustment for prematurity. Each test is standardized to have a mean of 100 and SD of 16 (B-I) and 15 (B-II). Mental development and psychomotor development indices were obtained at each testing occasion. When infants obtained raw scores too low to derive a scaled score, the lowest possible scaled score minus 1 (ie, 49) was imputed. This is a common technique for addressing floor effects but may result in overestimation of a child’s performance. Also, if the administering neuropsychologist indicated that the child was unable to complete the test under standardized test conditions, a value of 49 was imputed. Such tests were included because they have been found to be predictive of HIV disease progression.

Patient Populations

Three cohorts of perinatally exposed infants were identified. Cohort I included infants born between October 1992 and June 1997 (when the first PACTG study using a PI started enrollment) having ≥1 B-I evaluation before 1 year of age and was used to compare the profiles of HIV− and HIV+ infants before HAART with PIs became available. Any evaluations completed after a child started PIs were excluded. Cohort II included infants born after June 1997 having ≥1 B-II evaluation before 1
year of age. This cohort was used to describe and compare profiles of HIV+ (regardless of actual treatment regimen) and HIV− infants in the era of PI-based HAART and included follow-up until October 2005. Cohort III included infants having ≥1 B-II evaluation before and ≥1 after they started PIs (before or after 1 year of age and regardless of date of birth). Some infants in cohort III were also part of cohort II.

Sociodemographic, Health, and ART Covariates
Sociodemographic variables included age, gender, and race/ethnicity. Birth characteristics included birth weight, maternal history of intravenous drug use, and any maternal ART use. Primary caregiver and maximum years of schooling completed by the primary caregiver were used as surrogates for family status and socioeconomic status, respectively. CD4 lymphocyte percentage (<25% or ≥25%) and Centers for Disease Control and Prevention (CDC) disease category at the time of the first Bayley evaluation were used to characterize HIV health status. HIV-1 RNA levels were not collected in PACTG 219 before 2000 and, thus, were not included in this analysis. As a surrogate for general health, ratings (0%–100%) were calculated from 3 scales measuring overall health, physical, and emotional well-being. The number of stressful life events since the previous study visit (0–9: caregiver lost job, family member left home, loss of housing, entitlement or health insurance, previous study visit (0–9: caregiver lost job, family member left home, loss of housing, entitlement or health insurance, and Centers for Disease Control and Prevention (CDC) disease category at the time of the first Bayley evaluation were used to characterize HIV health status. HIV-1 RNA levels were not collected in PACTG 219 before 2000 and, thus, were not included in this analysis. As a surrogate for general health, ratings (0%–100%) were calculated from 3 scales measuring overall health, physical, and emotional well-being. The number of stressful life events since the previous study visit (0–9: caregiver lost job, family member left home, loss of housing, entitlement or health insurance, family member hospitalized or very sick, change in caregiver, death in the family) was used as a surrogate for general well-being.28 Using an intent-to-treat approach, subjects were classified as being on PI-based HAART after their first exposure to PIs, regardless of subsequent treatment changes.

Statistical Methods
The response variable was the Bayley scaled score, grouped into 6-month age intervals ≤24 months and at 36 months using the protocol-specified testing schedule: 6 (lower age limit of test to 9 months), 12 (>9 to 15 months), 18 (>15 to 21 months), 24 (>21 to 30 months), and 36 (>30 to 42 months). Fisher’s exact tests and t tests were used for univariate comparisons of categorical and continuous variables. Random-effects models fitting separate intercepts and slopes to each subject’s scaled scores were used to compare profiles between HIV− and HIV+ infants adjusting for other covariates.29 To test for changes in slopes before and after the initiation of PIs in cohort III, a random-effects model with a change point was used.30 Generalized estimating equations (GEEs) were used to compare proportions of HIV+ and HIV− infants with significant mental or motor delays (scaled scores <70) as they aged.31 To assess the influence of potentially informative missingness patterns, analyses were replicated using only observed results and then using the last observed value carried forward to missing time points. Because conclusions were similar, only results using the observed data are presented.

RESULTS
Patient Populations
Cohort I included 275 (221 HIV− and 54 HIV+) infants, and cohort II included 929 (838 HIV− and 91 HIV+) infants. There were a further 656 (558 HIV− and 98 HIV+) infants enrolled in PACTG 219/219C before 1 year of age who did not have the necessary B-I or B-II test before 1 year of age for inclusion in cohorts I or II. To assess whether those included in the 2 cohorts were different from those not included, we compared sociodemographic characteristics, birth characteristics, and health variables at entry to PACTG 219/219C. The variables that differed significantly were the proportion of mothers with a history of IV drug use (9% among those in cohort I/II vs 14%; P < .01) and, among the HIV+ subjects, higher CD4 percentage (means of 46% vs 31%; P = .03).

Cohort I: Cognitive Functioning in HIV− and HIV+ Infants and Young Children ≤2 Years of Age in the Pre-PI Era
Subject Characteristics
Characteristics of infants in cohort I at the time of their first B-I test are shown in Table 1. HIV− infants were younger than the HIV+ infants (mean of 6.7 months vs 7.6 months; P = .03). Their mothers had higher rates of ART use during pregnancy (79% vs 48% of known; P < .01) and lower rates of previous IV drug use (11% vs 25% of known; P = .02). HIV− infants were also more likely to be born after January 1, 1995, reflecting the increased use of ART and the success of zidovudine to reduce mother-to-child transmission.32 Significantly more HIV− subjects had birth weights ≥2500 g than HIV+ infants (91% vs 67%; P < .01). More HIV− infants were living with a biological parent (96% of known vs 83%; P < .01), and they had higher mean health ratings (89% vs 73%; P < .01) and CD4 percentages (46% vs 31%; P < .01).

Initial Evaluation
At the initial evaluation before 1 year of age, mean (SD) mental and motor scores for the HIV− infants were 105 (17) and 107 (16), respectively, as compared with 85 (22) and 77 (25) for the HIV+ infants (P < .01 for both; Table 2). Both mean scores were statistically significantly above the standard norm of 100 for HIV− infants, whereas both were significantly <100 for HIV+ infants (P < .01 for all). In addition, the HIV+ infants had statistically significantly higher percentages of floor (49) scores than the HIV− infants (11% vs 1% for mental scores [P < .01] and 22% vs 1% for motor scores [P < .01]).

Among HIV+ infants, mean motor scores were stas-
TABLE 1  Characteristics of Cohort I at Time of First B-I Evaluation

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>HIV− (n = 221)</th>
<th>HIV+ (n = 54)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>.10a</td>
</tr>
<tr>
<td>Male</td>
<td>98</td>
<td>44.3</td>
<td>31</td>
</tr>
<tr>
<td>Female</td>
<td>123</td>
<td>55.7</td>
<td>23</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td>.12a</td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>36</td>
<td>16.3</td>
<td>15</td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>90</td>
<td>40.7</td>
<td>24</td>
</tr>
<tr>
<td>Hispanic</td>
<td>89</td>
<td>40.3</td>
<td>14</td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
<td>2.7</td>
<td>1</td>
</tr>
<tr>
<td>English primary language?</td>
<td></td>
<td></td>
<td>&lt;.01a</td>
</tr>
<tr>
<td>Yes</td>
<td>150</td>
<td>67.9</td>
<td>48</td>
</tr>
<tr>
<td>Year of birth</td>
<td></td>
<td></td>
<td>&lt;.01a</td>
</tr>
<tr>
<td>Before Jan 1, 1995</td>
<td>128</td>
<td>57.9</td>
<td>49</td>
</tr>
<tr>
<td>On or after Jan 1, 1995</td>
<td>93</td>
<td>42.1</td>
<td>5</td>
</tr>
<tr>
<td>ART in pregnancy</td>
<td></td>
<td></td>
<td>&lt;.01a</td>
</tr>
<tr>
<td>Yes</td>
<td>175</td>
<td>79.2</td>
<td>25</td>
</tr>
<tr>
<td>No</td>
<td>46</td>
<td>20.8</td>
<td>27</td>
</tr>
<tr>
<td>Unknown</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Maternal history of intravenous drug use</td>
<td></td>
<td></td>
<td>.02a</td>
</tr>
<tr>
<td>Yes</td>
<td>25</td>
<td>11.4</td>
<td>13</td>
</tr>
<tr>
<td>No</td>
<td>194</td>
<td>88.6</td>
<td>39</td>
</tr>
<tr>
<td>Unknown</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Birth weight</td>
<td></td>
<td></td>
<td>&lt;.01a</td>
</tr>
<tr>
<td>&lt;2500 g</td>
<td>20</td>
<td>9.1</td>
<td>18</td>
</tr>
<tr>
<td>≥2500 g</td>
<td>199</td>
<td>90.9</td>
<td>36</td>
</tr>
<tr>
<td>Missing</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Primary caregiver</td>
<td></td>
<td></td>
<td>&lt;.01a</td>
</tr>
<tr>
<td>Biological parent</td>
<td>210</td>
<td>95.5</td>
<td>45</td>
</tr>
<tr>
<td>Other</td>
<td>10</td>
<td>4.6</td>
<td>9</td>
</tr>
<tr>
<td>Unknown</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Maximum education of primary caregiver</td>
<td></td>
<td></td>
<td>.81a</td>
</tr>
<tr>
<td>Primary</td>
<td>27</td>
<td>12.7</td>
<td>6</td>
</tr>
<tr>
<td>Secondary</td>
<td>124</td>
<td>58.5</td>
<td>28</td>
</tr>
<tr>
<td>Higher than secondary</td>
<td>61</td>
<td>28.8</td>
<td>17</td>
</tr>
<tr>
<td>Unknown</td>
<td>9</td>
<td>4.1</td>
<td>3</td>
</tr>
<tr>
<td>No. of stressful events</td>
<td></td>
<td></td>
<td>.61b</td>
</tr>
<tr>
<td>0</td>
<td>96</td>
<td>45.9</td>
<td>27</td>
</tr>
<tr>
<td>1–2</td>
<td>71</td>
<td>34.0</td>
<td>16</td>
</tr>
<tr>
<td>≥2</td>
<td>42</td>
<td>20.1</td>
<td>11</td>
</tr>
<tr>
<td>Unknown</td>
<td>12</td>
<td>5.5</td>
<td>0</td>
</tr>
<tr>
<td>CD4 percentage (%)</td>
<td></td>
<td></td>
<td>&lt;.01b</td>
</tr>
<tr>
<td>&lt;15%</td>
<td>1</td>
<td>0.5</td>
<td>6</td>
</tr>
<tr>
<td>15%–24%</td>
<td>2</td>
<td>1.0</td>
<td>10</td>
</tr>
<tr>
<td>≥25%</td>
<td>196</td>
<td>98.5</td>
<td>35</td>
</tr>
<tr>
<td>Missing</td>
<td>22</td>
<td>10.0</td>
<td>2</td>
</tr>
<tr>
<td>Health rating</td>
<td></td>
<td></td>
<td>&lt;.01b</td>
</tr>
<tr>
<td>≤85</td>
<td>52</td>
<td>25.7</td>
<td>32</td>
</tr>
<tr>
<td>&gt;85</td>
<td>150</td>
<td>74.3</td>
<td>21</td>
</tr>
<tr>
<td>Missing</td>
<td>19</td>
<td>9.0</td>
<td>1</td>
</tr>
</tbody>
</table>

All percentages are of subjects with known status.

* Data are from Fisher’s exact test, excluding missing.

b Data are from t test.

tically significantly higher among those whose mothers received ART during pregnancy (89 [23] vs 66 [20]; P < .01; Table 2). There was no significant difference in mental scores in the HIV+ infants by maternal ART use, although the trend was in the same direction as for motor scores (88 [23] vs 81 [22]; P = .26). The HIV− infants showed no differences in mental or motor scores by maternal ART use. Mental and motor scores were not statistically significantly different between infants (HIV+ or HIV−) whose mothers had a history of IV drug use.

Both mental (P = .04) and motor (P = .06) scores of HIV+ infants were positively correlated with CD4 percentages. Mean mental scores for infants with CD4 percentages <15%, 15% to 25%, and ≥25% were 66, 79, and 90, and mean motor scores were 53, 78, and 82. Mean mental scores were also associated with CDC disease category (A [93], B [79], and C [76]; P = .05). Mean motor scores by CDC disease category were 85, 73, and 69 (P = .17).

Comparison of Trajectories of B-I Scores ≤2 Years of Age

Data completeness at the upper age limit of the B-I is shown in Table 3. Up until the last B-I test, 7 (13%) of the HIV+ infants and young children had received no antiretroviral treatment, 45 (83%) had taken only nucleoside reverse transcriptase inhibitors (NRTIs) and 2 (4%) had taken NRTIs and non-NRTIs.

Figure 1A and B show mean (95% confidence intervals) scaled scores by age for the HIV− and HIV+ infants by CD4 percentage (<25% or ≥25%) at the time of the first B-I test. Numbers of tests in each age group and the percentage of mental floor scores are shown in Table 4. Percentages for the motor scores were similar and are not shown. Both HIV+ groups had significantly lower mental and motor scores than the HIV− infants at 6 months of age, and the observed differences remained relatively constant up to 24 months of age. The HIV− infants with CD4 percentages <25% had lower mean mental (P < .01) and motor (P < .01) scores than infants with CD4 percentages ≥25%. Motor scores were lower than mental scores over the entire age range. Random-effects models showed that both mean mental and motor scores declined significantly with age (−6.4 [SE: 1.1] points per year for mental scores and −3.5 [SE: 1.1] points per year for motor scores; P < .01 for both), but there was no evidence that the rate of decline differed between HIV− and HIV+ infants and young children (P = .40 and .55, respectively).

Cohort II: Cognitive Functioning in HIV− and HIV+ Infants and Young Children ≤3 Years of Age in the PI Era

Subject Characteristics

Subject characteristics of cohort II at the time of their first B-II test are shown in Table 5. More HIV− infants than HIV+ infants were born after 2001 (P = .04), and the mothers of HIV− infants had higher rates of ART use during pregnancy (P < .01), reflecting increasingly effective treatments to prevent mother-to-child transmission. The proportion of infants with birth weights ≥2500 g was significantly higher in HIV− than HIV+
infants (83% vs 67%; P < .01) and, as in cohort I, the HIV− infants had a lower rate of previous maternal IV drug use (6% vs 19%; P < .01). A higher proportion of HIV− infants was living with a biological parent (95% vs 76% of known; P < .01), and the HIV− infants had higher CD4 percentages than HIV+ infants (mean of 45% vs 37%; P < .01). Of the 86 HIV+ infants with complete ARV histories, 54 (63%) were on HAART with a PI, 8 (9%) were on HAART without a PI (2 NRTIs and 1 non-NRTI), and 24 (28%) were on other single or combination ARVs.

**Initial Evaluation**

Mean (SD) mental and motor scores for the HIV− infants were 92 (12) and 90 (15), respectively, as compared with 85 (16) and 83 (19) for the HIV+ infants (Table 6). Mean scores for both the HIV+ and HIV− infants were statistically significantly below the standard norm of 100. Mean mental and motor scores of HIV+ infants were also significantly lower relative to the HIV− infants (P < .01 for both). In addition, HIV+ infants had significantly higher percentages of floor scores (4% for HIV− vs 1% for HIV+ for mental tests [P < .01] and 6% vs 1% for motor scores [P = .02]). There were no statistically significant differences in mental or motor scores by maternal ART use during pregnancy or by history of IV drug use in either the HIV+ or HIV− infants. However, there was a trend for mean motor scores to be lower in the HIV+ infants whose mothers had not received ART dur-

| TABLE 2 | Mental and Motor Scores for Cohort I First Evaluation Before 1 Year of Age |
|----------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Characteristic      | Mental Score        | Motor Score         |                    |                    |                    |                    |                    |
|                     | N          | Mean (SD) | 95% CI | P  | N          | Mean (SD) | 95% CI | P  |
| HIV+                 |            |           |        |    |            |           |        |    |
| HIV−                 |            | 221       | 54     | 85 (22) | 79–91 | <.01*  | 221     | 54   | 77 (25) | 70–84 | <.01*  |
| Maternal ART exposure| HIV+       |            |        |    |            |           |        |    |
| Yes                  | 25         | 25        | 88 (23) | 79–98 | .26    | 25        | 89 (23) | 79–98 |
| No                   | 27         | 27        | 81 (22) | 73–90 |        | 27        | 66 (20) | 58–74 |
| HIV−                 |            |            | 175    | 105 (17) | 102–108 | .49    | 175     | 105 (16) | 105–110 |
| CD4 percentage       | <15%       | 6          | 66 (21) | 45–88 | .02b   | 6          | 53 (6)  | 47–59 |
|                       | 15%–24%    | 10         | 79 (21) | 64–94 |        | 10        | 78 (28) | 58–99 |
|                       | ≥25%       | 35         | 106 (17) | 83–97 |        | 194       | 107 (16) | 105–110 |
| CDC disease category | N or A     | 22         | 93 (18) | 85–101 | .04b   | 22        | 85 (29) | 72–97 |
|                       | B          | 13         | 79 (23) | 65–93 |        | 13        | 73 (20) | 61–85 |
|                       | C          | 15         | 76 (20) | 65–88 |        | 15        | 69 (23) | 56–82 |
|                       |            |            |        |        |        |        |        |    |
| TABLE 3 | Completeness of B-I and B-II Test Scores at Upper Age Limits |
| Variable                          | B-I HIV− | B-I HIV+ | B-II HIV− | B-II HIV+ |
| Test at upper age limit           | N | % | N | % | N | % | N | % |
| Complete                          | 135 | 61 | 32 | 59 | 179 | 21 | 32 | 35 |
| B-II/next age-appropriate test    | 25 | 11 | 6 | 11 | 87 | 10 | 10 | 11 |
| Lost to follow-up                 | 28 | 13 | 0 | 0 | 107 | 13 | 5 | 6 |
| Died                              | 0 | 0 | 8 | 15 | 2 | 0 | 1 | <1 |
| Not yet old enough                | 0 | 0 | 0 | 0 | 339 | 41 | 26 | 29 |
| Missing                           | 33 | 15 | 8 | 15 | 124 | 15 | 17 | 19 |
| Total                             | 221 | 100 | 54 | 100 | 838 | 100 | 91 | 100 |

CI indicates confidence interval. 
* Data are from t-test. 
** Data are from linear regression.
ing pregnancy (80 [SD: 22] vs 87 [SD: 14]; P = .06).
Among HIV+ infants, neither mental nor motor scores were correlated with CD4 percentages or CDC disease category.

Comparison of Trajectories of B-II Scaled Scores ≤3 Years of Age
A summary of data completeness at the upper age limit of the B-II is shown in Table 3. Among the 86 HIV+ infants and young children with complete ARV history, the percentage on PIs increased from 58% (n = 52) at the 6-month test to 77% (n = 31) at the 36-month test. The percentage of subjects not on HAART decreased from 35% at 6 months to 19% at 36 months.

Figure 2A and B and Table 6 show mean (95% confidence intervals) scaled scores by age and HIV status. Numbers of tests in each age group and the percentage of mental floor scores are shown in Table 4. At 6 months, the HIV+ infants had significantly lower mean mental and motor scores than the HIV- infants regardless of immunologic status (CD4 percentage < or ≥25%). By 24 and 36 months, these differences were no longer statistically significant. Random-effects models showed that mental scores declined at a significantly greater rate in the HIV- infants (−6.2 [SE: 0.4] points per year) than the HIV+ infants (−3.2 [SE: 1.0] points per year; P = .01). The same pattern was observed for motor scores, with the HIV- infants and young children changing by an average of −1.4 (SE: 0.4) points per year, whereas the HIV+ infants and young children showed improvements of 1.3 (SE: 1.2) points per year (P = .03). Models were then fit including 2-way interactions of the study under which the infant had enrolled (219 or 219C), each sociodemographic, birth characteristic, and health covariate with age. Although the statistical significance of the effect of HIV status on the rates of change varied, the direction was consistent (HIV+ infants improved relative to HIV- infants over time), and the effect remained at least marginally significant (P ≤ .07). As with mental score models, the trend for motor scores by HIV status was consistently for the HIV+ infants and young children to improve relative to the HIV- infants and young children with statistical significance for this effect ranging from P = .04 to P = .15. There was no evidence that any of these factors were affecting the rates of change differently in the HIV+ or HIV- infants.

The percentage of infants and young children in cohort II with significant mental and motor delays (scores ≤70 including floor scores) was significantly higher among HIV+ than HIV- infants and young children at all ages (Table 7). However, the difference decreased with age (P < .01, GEE model), reflecting a greater increase in the proportion of HIV+ infants and young children with significant impairment with age than observed in HIV+ infants and young children.

Cohort III: Impact of PIs on B-II Scores of HIV+ Infants and Young Children
Ninety-one children had a B-II test score before starting PIs, with ≥1 subsequent B-II test after the initiation of PIs and before 42 months of age, 13 of whom were also in cohort II. The median age at initiation of PIs in this group was 20.4 months (10th and 90th percentiles: 9.6 and 33.6 months, respectively). More than half (n = 49 or 54%) were born before 1996 and had no opportunity
to start PIs before 1 year of age, 29 (32%) were born in 1996–1997, and the remaining 19 (14%) were born after 1997 when PIs were more widely available.

Random-effects models were fit to compare slopes in mental and motor scores before and after PIs. Before starting a PI, mental scores declined an average of −8.4 (SE: 1.8) points per year (P < .01; reference rate of decline in HIV− infants was −6.2 [SE: 0.4] points per year from cohort II), and after starting PIs, this rate of decline was significantly reduced to −1.1 (SE: 2.2) points per year (P = .02). A similar trend for motor scores was observed. The average rate of decline before starting a PI was −3.1 (SE: 2.1) points per year, and after starting, there was a mean increase of +0.7 (SE: 2.5) points per year, but the difference was not significant (P = .26). Sensitivity analyses using the sociodemographic and health covariates did not influence these findings.

**DISCUSSION**

The use of HAART for HIV+ children has reduced mortality and morbidity during recent years. The efficacy of HAART with respect to distal outcomes, such as neurodevelopmental functioning, has not been well described, despite the fact that HAART is now a standard of care in pediatric HIV infection. The results of this investigation are consistent with investigations during the pre-HAART era that identified compromised mental and motor functioning among infants with HIV infection. The results also extend our understanding of the effects of PI-based HAART showing a positive but limited impact on neurodevelopmental trajectories and the rate of significant cognitive and motor impairment in HIV+ infants and young children during the first 3 years of life, despite substantial effects on survival and immunologic status. This limited positive effect of HAART may be partly because infants and young children with compromised CNS now survive the more severe early manifestations of the disease but may have persistent neurobehavioral difficulties.

The observed differences in neurodevelopment between HIV+ and HIV− infants in cohort I may be related to genetic, health, disease, treatment, and/or psychosocial factors. For example, HIV− infants were significantly more likely to have been exposed to ART in utero and to weigh more at birth. They were less likely to have mothers with a history of IV drug use and more likely to live with a biological parent. HIV+ infants, on the other hand, had lower mean health ratings and CD4 percentages, notwithstanding ART that may have been provided during infancy. Infants with immune suppression (CD4 percentages <25%) seemed to be most vulnerable with respect to neurodevelopmental functioning, demonstrating mental and motor scores lower than those of the HIV− infants, as well as HIV+ infants with less immune suppression (CD4 percentages ≥25%). As Tardieu suggests, these very young children may experience a specific form of HIV-associated CNS disease related to prenatal HIV infection of the brain. Similar results have been shown in a recent study comparing McCarthy Scales of Children’s Abilities neurodevelopmental scores for HIV+ children with and without an early AIDS-defining illness and HIV-exposed but uninfected children aged 3 to 7 years. It is worth noting that the difference between HIV+ and HIV− infants was already present in all of the cohorts at the earliest measurement point, before 1 year of age. This finding is similar to the results from Llorente et al, who found significant developmental differences at 4 months of age that were predictive of survival and also illustrated the significant neurobehavioral effects of HIV in the prenatal and early postnatal period.

Both HIV− and HIV+ infants and young children in cohort I displayed a negative developmental trajectory, with indices declining at similar rates from birth to 2 years of age. These results may reflect the relative det-

**TABLE 4** Number and Percentage of Mental Floor Scores by Age Group

<table>
<thead>
<tr>
<th>Variable</th>
<th>HIV− CD4 Percentage &lt;25%</th>
<th>HIV− CD4 Percentage ≥25%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total N</td>
<td>Floor</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Month 6</td>
<td>171</td>
<td>4</td>
</tr>
<tr>
<td>Month 12</td>
<td>159</td>
<td>1</td>
</tr>
<tr>
<td>Month 18</td>
<td>128</td>
<td>0</td>
</tr>
<tr>
<td>Month 24</td>
<td>135</td>
<td>1</td>
</tr>
<tr>
<td>Month 6</td>
<td>593</td>
<td>4</td>
</tr>
<tr>
<td>Month 12</td>
<td>581</td>
<td>4</td>
</tr>
<tr>
<td>Month 18</td>
<td>283</td>
<td>4</td>
</tr>
<tr>
<td>Month 24</td>
<td>350</td>
<td>23</td>
</tr>
<tr>
<td>Month 30</td>
<td>179</td>
<td>10</td>
</tr>
</tbody>
</table>

The column labeled “Raw N” shows the number of tests assigned a floor score because raw scores were too low to scale.
rimental impact of genetic susceptibility\textsuperscript{38} and a psychosocial environment often characterized by poverty and other risk factors known to influence cognitive development\textsuperscript{39,40} and often observed in HIV-affected families. Declines of similar magnitude on the Bayley over the first years of life have been reported for non–HIV-affected groups from comparable socioeconomic backgrounds\textsuperscript{41,42}.

HAART has a substantial impact on the immunologic status of HIV\textsuperscript{+} children\textsuperscript{13}. In our study, HIV\textsuperscript{+} infants in cohort II had higher mean CD4 percentages at 6 months relative to those in cohort I and their CD4 percentages

\begin{table}[h]
\centering
\caption{Characteristics of Cohort II at Time of First B-II Evaluation}
\begin{tabular}{l|cc|cc|c}
\hline
\textbf{Characteristic} & \multicolumn{2}{c}{\textbf{HIV\textsuperscript{+} (n = 838)}} & \multicolumn{2}{c}{\textbf{HIV\textsuperscript{-} (n = 91)}} & \textbf{P} \\
& \textbf{N} & \textbf{%} & \textbf{N} & \textbf{%} & \\
\hline
\textbf{Gender} & & & & & .74\textsuperscript{a} \\
\text{Male} & 421 & 50.2 & 44 & 48.4 & \\
\text{Female} & 417 & 49.8 & 47 & 51.6 & \\
\hline
\textbf{Race/ethnicity} & & & & & .56\textsuperscript{a} \\
\text{White non-Hispanic} & 77 & 9.2 & 6 & 6.6 & \\
\text{Black non-Hispanic} & 511 & 61.0 & 56 & 61.5 & \\
\text{Hispanic} & 236 & 28.2 & 26 & 28.6 & \\
\text{Other} & 14 & 1.7 & 3 & 3.3 & \\
\hline
\textbf{English primary language} & & & & & .51\textsuperscript{a} \\
\text{Yes} & 644 & 76.8 & 73 & 80.2 & \\
\text{No} & 173 & 20.4 & 28 & 31.5 & \\
\hline
\textbf{Year of birth} & & & & & .04\textsuperscript{a} \\
<2001 & 288 & 34.4 & 43 & 47.3 & \\
2001–2002 & 256 & 30.5 & 26 & 28.6 & \\
>2002 & 294 & 35.1 & 22 & 24.2 & \\
\hline
\textbf{ART in pregnancy} & & & & & <.01\textsuperscript{a} \\
\text{Yes} & 793 & 96.6 & 40 & 46.5 & \\
\text{No} & 28 & 3.4 & 46 & 53.5 & \\
\text{Unknown} & 17 & 4 & 3 & 3.3 & \\
\hline
\textbf{Maternal history of intravenous drug use} & & & & & <.01\textsuperscript{a} \\
\text{Yes} & 42 & 5.5 & 15 & 19.2 & \\
\text{No} & 726 & 94.5 & 63 & 80.8 & \\
\text{Unknown} & 70 & 9 & 2 & 2.5 & \\
\hline
\textbf{Birth weight} & & & & & <.01\textsuperscript{a} \\
<2500 g & 144 & 17.2 & 30 & 33.3 & \\
\geq 2500 g & 692 & 82.8 & 60 & 66.7 & \\
\text{Missing} & 2 & 1 & 1 & 1 & \\
\hline
\textbf{Primary caregiver} & & & & & <.01\textsuperscript{a} \\
\text{Biological parent} & 793 & 94.9 & 66 & 75.9 & \\
\text{Other} & 43 & 5.1 & 21 & 24.1 & \\
\text{Unknown} & 2 & 4 & 1 & 1 & \\
\hline
\textbf{Maximum education of primary caregiver} & & & & & .55\textsuperscript{a} \\
\text{Primary} & 64 & 8.2 & 5 & 6.1 & \\
\text{Secondary} & 462 & 58.9 & 45 & 54.9 & \\
\text{Higher than secondary} & 259 & 32.9 & 32 & 39.0 & \\
\text{Unknown} & 53 & 9 & 5 & 5.5 & \\
\hline
\textbf{No. of stressful events} & & & & & .98\textsuperscript{b} \\
0 & 581 & 71.4 & 56 & 68.3 & \\
1–2 & 157 & 19.3 & 21 & 25.6 & \\
>2 & 76 & 9.3 & 5 & 6.1 & \\
\text{Unknown} & 24 & 3 & 1 & 1 & \\
\hline
\textbf{CD4 percentage} & & & & & <.01\textsuperscript{b} \\
<15\% & 0 & 0.0 & 3 & 3.7 & \\
15\%–24\% & 4 & 0.6 & 6 & 7.4 & \\
\geq 25\% & 713 & 99.4 & 72 & 88.9 & \\
\text{Missing} & 121 & 10 & 10 & 10 & \\
\hline
\textbf{Health rating} & & & & & .14\textsuperscript{b} \\
\leq 85 & 134 & 18.3 & 16 & 20.5 & \\
>85 & 599 & 81.7 & 62 & 79.5 & \\
\text{Missing} & 105 & 13 & 10 & 10 & \\
\hline
\end{tabular}
\end{table}

All of the percentages are of subjects with known status.
\textsuperscript{a} Data are from Fisher’s exact test, excluding missing.
\textsuperscript{b} Data are from \textit{t} test.

\textsuperscript{38} LINDSEY et al

\textsuperscript{13} at Prince Of Songkla Univ on April 19, 2007 www.pediatrics.orgDownloaded from
improved relative to the natural rate of decline in HIV− infants and young children ≤3 years of age (data not shown). This group also showed limited but positive trends in neurodevelopmental functioning relative to their HIV− peers, both in analyses looking at changes in test scores and in the proportion of infants with severe impairment, about which we may be cautiously optimistic. Initial differences in mental functioning between HIV+/H11002 and HIV−/H11002 infants in the post-PI era (cohort II) diminished with age, reflecting the more positive trajectory in scores in HIV−/H11001 versus HIV+/H11002 infants and young children before and after starting a PI in cohort III. The trends over age in cohort II were observed in the context of gradually declining mental and motor scores in HIV+/H11002 infants and young children, as also observed among infants in cohort I. Again, the impact of poverty and other psychosocial factors on neurodevelopmental functioning is powerful and must continue to be considered as we attempt to understand the role of HAART in the health and development of children with HIV infection.

There are a number of limitations to this analysis, including the confounding in calendar time of the transition from the B-I to B-II in the parent study, which occurred as PIs were introduced in children. “Transforming” scores from 1 neurodevelopmental test onto an updated version is not feasible in this population because many infants score well below the normed mean of 100, and there are no published data on the differences in test scores at the lower end of the scale. We cannot, therefore, directly compare scores in cohorts I and II. However, we believe the within-subject changes in scores over time may be comparable in the 2 test versions. In the HIV+/H11002 infants and young children, the rate of decline in mental scores was −6.4 (SE 1.1) points per year on the B-I and −6.2 (0.4) points per year on the B-II, but only in cohort II were the rates of decline smaller in the HIV−/H11001 infants and young children, providing evidence that neurodevelopmental status was better in the post-PI era. In addition, the scores obtained by infants on the B-I, published in 1969, may be subject to the Flynn effect, which is the systematic increase in intelligence quotient scores that causes test norms to become obsolete over time.43,44 B-II results likely provide a more valid assessment of the developmental status of infants and young children in cohort II, because they are based on recently updated norms from a contemporary sample.

With the small number of HIV+/H11001 infants in cohort III, we used a simple classification of starting PI-based HAART, not allowing for variability in treatment regimens and switches over time. This “intent-to-treat” approach is not unrealistic, however, because once infants start PIs, in general they stay on PIs, albeit with occasional interruptions. Our ability to observe the true ef-

<table>
<thead>
<tr>
<th>TABLE 6</th>
<th>Mental and Motor Scores for Cohort II First Evaluation Before 1 Year of Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>Mental Score</td>
</tr>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>HIV+</td>
<td>91</td>
</tr>
<tr>
<td>HIV−</td>
<td>838</td>
</tr>
<tr>
<td>Maternal ART exposure HIV+</td>
<td>40</td>
</tr>
<tr>
<td>HIV-</td>
<td>26*</td>
</tr>
<tr>
<td>Maternal history of intravenous drug use HIV+</td>
<td>793</td>
</tr>
<tr>
<td>HIV−</td>
<td>49*</td>
</tr>
<tr>
<td>CD4 percentage</td>
<td></td>
</tr>
<tr>
<td>&lt;15%</td>
<td>3</td>
</tr>
<tr>
<td>15%–24%</td>
<td>6</td>
</tr>
<tr>
<td>≥25%</td>
<td>72</td>
</tr>
<tr>
<td>CDC disease category</td>
<td></td>
</tr>
<tr>
<td>N or A</td>
<td>47</td>
</tr>
<tr>
<td>B</td>
<td>6</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
</tr>
</tbody>
</table>

CI indicates confidence interval.
* Data are from t test.
** Data are from linear regression.
effects of PIs on cognitive scores may also have been reduced because of confounding by indication, in which newer treatments are accessed first by children with more advanced disease and reduced potential for normal development, thus underestimating the effect of starting a PI-based HAART.

Another potential limitation concerns selection effects if test scores were more likely to be missing because the infants were neurologically impaired. To enter the analysis cohorts, neurodevelopmental evaluation with a Bayley scale was required at ≤1 year of age. If impaired infants were more likely not to have a test administered, the cohorts would not be representative of the broader PACTG 219/219C population. There was some evidence that this was the case, because those in cohorts I and II had higher mean CD4 percentages and a lower proportion of mothers with a history of intravenous drug use than those not evaluated. There were also missing test results at the upper age limits of the scales. Reasons included loss to follow-up from the study, not having a test in the required window, and transitioning to the next age-appropriate test. If HIV+ infants with more advanced impairment or HIV− infants with normal functioning were more likely to be lost to follow-up, miss visits, or be more or less likely to be evaluated with the test aimed at younger infants, then the trajectories we observed over the first 3 years are likely to be biased. There was some evidence that the higher-functioning infants and young children were more likely to move on to the next age-appropriate test at the test transition ages (results not shown). However, we repeated the analyses imputing the last nonmissing score at each testing window for infants lost to follow-up (not including those who transitioned to the next age-appropriate test), and qualitatively the results did not change. Additional but unmeasurable factors, such as use of early intervention services and the effects of repeated testing using the same instrument, were not controlled for and may have affected the results. However, practice effects may be less likely when developmental testing is completed with infants and young children, because item sets vary according to the age of the child.

Finally, the random-effects models were fit using the imputed value of 49 for the floor scores, potentially violating the assumptions of normality. Because the results from the GEE analyses were consistent with the analyses using the continuous data, we do not believe that the imputed scores greatly influenced the results.

CONCLUSIONS
The etiology and natural history of neurodevelopmental impairments associated with HIV infection in young children remain difficult to identify and predict, because neurodevelopment in these children is determined by the interaction of multiple genetic, health, disease, treatment, and psychosocial factors that may vary in their impact during infancy and childhood. Further analyses focusing on the impact of maternal risk factors and ARV use with this cohort would be of interest. However, this study provides evidence of limited beneficial effects of PI-based HAART on neurodevelopment in the first years of life. The effects of HAART may become more apparent as the children reach school age and beyond and as differential and characteristic patterns of strengths and

![Figure 2](https://example.com/figure2.png)

**A**, Cohort II; mean B-II mental scaled scores according to age. **B**, Cohort II; mean B-II motor scaled scores according to age. The bars show 95% confidence intervals. The dashed lines show the HIV+ cohort divided into 2 groups, defined by the infants’ CD4 percentages measured at the time of first test before 1 year of age.
weakness in domains of development are identified. Results observed in our investigation necessarily require confirmation and elaboration in other studies.

**ACKNOWLEDGMENTS**

This work was supported by the Statistical and Data Analysis Center of the PACTG at the Harvard School of Public Health under the National Institute of Allergy and Infectious Diseases cooperative agreement 5U01 AI41110, the National Institute of Child Health and Human Development, and the National Institute of Mental Health.


**TABLE 7**

Mental and Motor B-II Scaled Scores at 6, 24, and 36 Months of Age

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mental Scores</th>
<th>Motor Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HIV+ Mean (SD) [95% CI] N [% &lt;70]</td>
<td>HIV+ Mean (SD) [95% CI] N [% &lt;70]</td>
</tr>
<tr>
<td>Age, mo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>92 (12) [91–93] 593 (5.2) 84 (16) [80–88] 57 (25.0)</td>
<td>&lt;0.01* 89 (15) [87–90] 593 (10.3) 80 (17) [76–85] 57 (35.4)</td>
</tr>
<tr>
<td>24</td>
<td>79 (17) [78–81] 350 (24.0) 75 (17) [70–80] 47 (41.8)</td>
<td>1.2* 87 (17) [86–89] 349 (16.6) 84 (17) [79–89] 47 (37.3)</td>
</tr>
<tr>
<td>36</td>
<td>80 (16) [77–82] 179 (24.6) 78 (14) [73–83] 32 (30.5)</td>
<td>4.6* 86 (18) [83–88] 179 (21.8) 84 (15) [79–83] 32 (30.5)</td>
</tr>
</tbody>
</table>

Cl indicates confidence interval.

a Data are from t test, P comparing HIV- with HIV+.

b Data are from Wald test for interaction of HIV status and age from random-effects model.

@ indicates confidence interval.

6.2 (0.4) .01 b

3.2 (1.0) .01 b
Eddleman; Children’s Medical Center of Dallas; Connecticut Children’s Medical Center: J. Salazar, G. Karas, L. Wells, and T. George; University of South Alabama: M. Mancuso; University of Chicago Children’s Hospital; Cook County Hospital: J. B. McAuley, K. M. Boyer, M. Haak, and N. Mourikis; Children’s Hospital at Albany Medical Center: A. D. Fernandez, P. A. Hughes, N. Wade, and M. E. Adams; State University of New York Upstate Medical University: L. B. Weiner, K. A. Contello, W. A. Holz, and M. J. Famiglietti; University of Florida at Gainesville: R. M. Lawrence, J. F. Lew, C. Delaney, and C. Duff; University of Rochester Medical Center: G. A. Weinberg, F. Gigliotti, B. Murante, and S. Laverty; Mount Sinai Medical Center: D. Johnson, D. Kowalski, and B. Wolfe; Public Health Unit of Palm Beach County: J. Sleasman and C. Delaney; Children’s Hospital and Medical Center, Seattle; Yale University School of Medicine: W. A. Andiman, S. Romano, L. Hurst, and D. Schroeder; St Josephs Hospital and Medical Center, Patterson: N. Hutchcon and A. Townley; Harbor-University of California Los Angeles Medical Center: M. Keller, C. Mink, S. Wettgen, and N. Redjal; Long Beach Memorial: A. Deveikis, L. Melton, S. Marks, and K. Elkins; Children’s Hospital of Los Angeles: J. Church and T. Dunaway; Lincoln Medical and Mental Health Center; Medical College of Georgia: C. S. Mani; Phoenix Children’s Hospital: J. P. Piatt, J. Foti, and L. Clarke-Steffen; Robert Wood Johnson Medical School: S. Gaur, P. Whitley-Williams, A. Malhotra, and L. Cerracchio; Vanderbilt University Medical Center: G. Wilson; University of Mississippi Medical Center: H. Gay and S. Sadler; Emory University Hospital: S. Nesheim and R. Dennis; Columbus Children’s Hospital: M. Brady, J. Hunkler, and K. Koranyi; University of South Florida: P. Emmanuel, J. Lujan-Zilberman, C. Graisberry, and S. Moore; Children’s Hospital of the King’s Daughters: R. G. Fisher, V. Van de Water, T. T. Rubio, and D. Sandler; Incarnation Children’s Center, New York: A. Gershon and P. Miller; Medical University of South Carolina: G. M. Johnson and A. Hutto; San Francisco General Hospital: D. Wara, A. Kamrin, and S. Farrales; Mt Sinai Children’s Hospital: M. Dolan, J. D’Agostino, and R. Posada; Beth Israel Medical Center; Medical College of Virginia: S. R. Lavoie and T. Y. Smith; The Medical Center, Pediatric, Columbus: C. Mani and S. Cobb; Children’s Hospital of Los Angeles: A. Kovacs and E. Operskalski; Cooper Hospital-University Medical Center: A. Feingold and S. Burrows-Clark; Sacred Heart Children’s Medical Services of Florida: W. Albritton; St Luke’s/Roosevelt Hospital Center: R. Warford and S. Arpadi; University of Cincinnati: J. Mrus, R. Beiting, and N. Boosveld; North Shore University Hospital: S. Pahwa and L. Rodriguez; Westchester Hospital; Metropolitan Hospital; Montefiore Medical-Albert Einstein College of Medicine; A. Rubinstein and G. Krikenik; Ceder’s/Sinai Medical Center; Case Western/Rainbow Babies & Children’s Hospital; Hermann Hospital; Lincoln Hospital; King’s County Hospital Center; St Louis Children’s Hospital: K. A. McGann, L. Pickering, and G. A. Storch; Georgetown University Hospital; Oregon Health and Science University: P. Lewis and R. Croteau.

We thank the children and their families, the study team, and the individuals and institutions involved in the conduct of PACTG 219C for contributing to this research.

REFERENCES


Infectious Disease Morbidity Among Young HIV-1–Exposed But Uninfected Infants in Latin American and Caribbean Countries: The National Institute of Child Health and Human Development International Site Development Initiative Perinatal Study

Marisa M. Mussi-Pinhata, MD\textsuperscript{a}, Laura Freimanis, MD, PhD\textsuperscript{b}, Aparecida Y. Yamamoto, MD\textsuperscript{a}, James Korelitz, PhD\textsuperscript{b}, Jorge A. Pinto, MD\textsuperscript{a}, Maria L. S. Cruz, MD\textsuperscript{b}, Marcelo H. Losso, MD\textsuperscript{b}, Jennifer S. Read, MD\textsuperscript{f}, for the National Institute of Child Health and Human Development International Site Development Initiative Perinatal Study Group

\textsuperscript{a}University of Sao Paulo, Ribeirão Preto, Sao Paulo, Brazil; \textsuperscript{b}Westat, Rockville, Maryland; \textsuperscript{c}Federal University of Minas Gerais, Belo Horizonte, Brazil; \textsuperscript{d}Hospital dos Servidores do Estado, Rio de Janeiro, Brazil; \textsuperscript{e}Hospital General de Aguadó Jose María Ramos Mejía, Buenos Aires, Argentina; \textsuperscript{f}Pediatric, Adolescent, and Maternal AIDS Branch, Center for Research for Mothers and Children, National Institute of Child Health and Human Development, Bethesda, Maryland

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

ABSTRACT

OBJECTIVE. The goal was to describe the frequency, characteristics, and correlates of infectious disease morbidity during the first 6 months of life among HIV-1–exposed but uninfected infants.

METHODS. The study population consisted of infants enrolled in the National Institute of Child Health and Human Development International Site Development Initiative Perinatal Study who were HIV-1 uninfected and had follow-up data through the 6-month study visit. Definitive and presumed infections were recorded at study visits (birth, 6–12 weeks, and 6 months).

RESULTS. Of 462 HIV-1–uninfected infants with 11,644 child-weeks of observation, 283 experienced ≥1 infection. These 283 infants experienced 522 infections (1.8 infections per infant). The overall incidence rate of infections was 4.5 cases per 100 child-weeks of observation. Overall, the most common infections were skin or mucous membrane infections (1.9 cases per 100 child-weeks) and respiratory tract infections (1.7 cases per 100 child-weeks). Thirty-six percent of infants had ≥1 respiratory tract infection (1.8 cases per 100 child-weeks). Incidence rates of upper and lower respiratory tract infections were similar (0.89 cases per 100 child-weeks and 0.9 cases per 100 child-weeks, respectively). Cutaneous and/or oral candidiasis occurred in 48 neonates (10.3%) and 92 older infants (19.3%). Early neonatal sepsis was diagnosed in 12 infants (26.0 cases per 1000 infants). Overall, 81 of 462 (17.5%) infants were hospitalized with an infection. Infants with lower respiratory tract infections were hospitalized frequently (40.7%). The occurrence of ≥1 neonatal infection was associated with more-advanced maternal HIV-1 disease.

Key Words
HIV-1, infancy, infections, Latin America, Caribbean

Abbreviations
NICHD—National Institute of Child Health and Human Development
NISDI—NICHD International Site Development Initiative
OR—odds ratio
CI—confidence interval

Accepted for publication Sep 25, 2006
Address correspondence to Marisa M. Mussi-Pinhata, MD, Departamento de Puericultura e Pediatria da FMRP-USP, Avenida Bandeirantes 3900, 14849-900 Ribeirão Preto, SP, Brazil. E-mail: mmmpinha@fmrp.usp.br

PEDIATRICS (ISSN Numbers: Print, 0031–4005; Online, 1098–4275). Copyright © 2007 by the American Academy of Pediatrics
tobacco use during pregnancy, infant anemia, and crowding. Lower maternal CD4+ cell counts, receipt of intrapartum antibiotic treatment, and country of residence were associated with postneonatal infections.

CONCLUSIONS. Close monitoring of HIV-1–exposed infants, especially those who are anemic at birth or whose mothers have more-advanced HIV-1 disease or who smoked during pregnancy, remains important.

Infectious diseases account for nearly 90% of early childhood deaths in developing countries. A substantial proportion of this morbidity and death occurs among young infants. Protection against viral and bacterial infections during early infancy (when innate and specific host defenses are maturing) is conferred primarily through breastfeeding and placental transfer of specific antibodies.

Because current prophylaxis regimens have been successful in decreasing mother-to-child transmission of HIV-1, uninfected infants born to HIV-1–infected mothers represent a growing population. In developing countries, such infants are vulnerable to acquiring infectious diseases if they are not themselves HIV-1 infected. Poor sanitation, limited maternal education, avoidance of breastfeeding, deficient transplacental transfer of IgG antibodies, and household crowding or close contact with immunodeficient individuals colonized with diverse pathogens may play a role. However, little information is available regarding patterns of infectious disease morbidity among HIV-1–exposed infants. To our knowledge, such information for infants from Latin America and the Caribbean region is not available.

The National Institute of Child Health and Human Development (NICHD) International Site Development Initiative (NISDI) Perinatal Study, a prospective cohort study of HIV-1–infected mothers and their infants in Latin America and the Caribbean region, represents a unique opportunity for analyzing the pattern of infectious morbidity among infants in these countries. The objective of this analysis was to describe the frequency, characteristics, and correlates of infectious disease morbidity occurring during the first 6 months of life among HIV-1–exposed infants.

METHODS

NISDI Perinatal Study Protocol

The primary objectives of the NISDI Perinatal Study include characterizing adverse events during pregnancy, the postpartum period, and early infancy. As requirements for enrollment of HIV-1–infected women into this protocol, antiretroviral prophylaxis and alternatives to breast milk needed to be available. Enrollment began in September 2002, and is ongoing. All subjects provide signed informed consent before enrollment into the study. The protocol was approved by the ethical review board at each clinical site where subjects were enrolled, as well as by institutional review boards at the sponsoring institution (NICHD) and the data coordinating center (Westat).

HIV-1–infected women are enrolled in the study during pregnancy and are monitored until 6 months after delivery. A medical history is obtained and a physical examination is conducted at each visit. Laboratory studies are conducted at most visits. The infants are examined at 3 study visits, namely, before hospital discharge after birth and at 6 to 12 weeks and 6 months of age. A medical history is obtained, a physical examination is performed (with assessments of growth and morbidity), and laboratory studies (flow cytometry studies and HIV-1 diagnostic assays, hematologic assays, and biochemical assays) are performed at each study visit. As part of the interval history, information is collected regarding the infant’s symptoms, diagnoses, medications, and hospitalizations.

When infants became ill, mothers were advised to seek medical attention at the clinical site; if medical attention was sought, then clinical, radiographic, and/or laboratory documentation would be available at the site. Therefore, in addition to data collected at NISDI Perinatal Study visits, clinical, radiographic, and/or laboratory data collected as part of routine medical care at the clinical site (and included in the subject’s medical charts) could be obtained. However, for mild events such as upper respiratory tract infections (eg, rhinitis or conjunctivitis) or mild diarrhea for which the mother did not seek medical attention, there would be no clinical, radiographic, or laboratory data available. Infants were considered HIV-1 uninfected if (1) they had ≥2 negative HIV-1 virologic assays (eg, HIV-1 culture or HIV-1 DNA polymerase chain reaction assays), with 1 test performed at ≥1 month of age and 1 performed at ≥4 months of age, and no positive virologic tests; (2) they had 1 positive HIV-1 virologic assay and 2 later HIV-1 virologic tests were negative (≥1 of which was performed at ≥4 months of age); or (3) they had 2 negative HIV-1 antibody tests (≥1 of which was performed at ≥6 months of age).

Study Population and Definitions

The study population for this analysis included infants enrolled in the NISDI Perinatal Study as of October 1, 2004, who completed follow-up monitoring in the study until 6 months of age and who were known to be HIV-1 uninfected. An infection was defined as having occurred when a definitive or presumptive diagnosis of infection was recorded. Specific standardized guidelines regarding the recording of diagnoses were followed by the sites. Definitive diagnosis required, in the majority of cases, that an organism be identified or serologic and/or antigenic evidence be found, unless the clinical findings
revealed the causative agent (eg, varicella zoster infection). Otherwise, the diagnosis was classified as presumed. An infection was considered incident if the infant had been free of any symptoms consistent with infection for the previous week.

Infections were categorized as congenital, systemic, central nervous system, upper respiratory tract, lower respiratory tract, gastrointestinal, renal and urinary system, or skin and mucous membranes. Incident infections were tallied at the date of initial diagnosis. Specific diagnoses included in each category, as well as the protocol definitions for each, are listed in the Appendix. Collection of microbiologic specimens was at the discretion of the subject’s physician and was not mandated by the study protocol. Because of the possible overlap in the clinical diagnoses of bronchiolitis and bronchitis in this age group, the 2 categories were combined for analysis as bronchiolitis.

Infant gestational age at birth (in completed weeks) was determined through either obstetrical estimation or pediatric newborn examination. Centers for Disease Control and Prevention growth standard reference curves were used for categorization of each subject’s weight for age at birth and postnatally. Infant hemoglobin values and absolute neutrophil counts and maternal CD4+ cell counts were classified as normal or abnormal (grade ≥1) according to the Division of AIDS grading system. Infant absolute lymphocyte counts were classified as normal or abnormal by using ranges of normal laboratory values appropriate for age. Maternal clinical disease staging was performed with the use of the 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. If antiretroviral drugs were received during pregnancy, then the reason for receipt of antiretroviral drugs was categorized as prophylaxis or treatment (prophylaxis: the patient was not receiving antiretroviral drugs when she became pregnant but administration of ≥1 antiretroviral drug was initiated during pregnancy and discontinued at or before the 6-12-week postpartum visit; treatment: the patient was receiving antiretroviral drugs before pregnancy and/or continued antiretroviral drug therapy after the 6-12-week postpartum visit).

Statistical Analyses

Frequencies of infections were calculated according to the age at the initial diagnosis, that is, early neonatal (0–6 days), late neonatal (7–27 days), or postneonatal (≥28 days). Incidence rates were calculated by using person-weeks as the denominator. Proportions were calculated and 95% confidence intervals (CIs) were calculated by using the exact binomial method. Associations of categorical variables with neonatal and postneonatal infections were evaluated by using the Fisher-Freeman-Halton exact test. Variables at least marginally associated with outcome (P ≤ .20) were considered candidates for the multivariate logistic regression modeling. Both stepwise selection and backward elimination strategies were applied, to determine whether the 2 selection procedures arrived at the same parsimonious model (using a 5% significance level).

RESULTS

Derivation of the Study Population

As of October 2004, 700 infants had been born to HIV-1–infected women enrolled in the NISDI Perinatal Study. Of those 700 infants, 487 had completed follow-up monitoring through 6 months of age, 204 were still undergoing follow-up monitoring in the study, 6 had died, and 3 had been lost to follow-up monitoring. Of the 487 infants who had completed the protocol, 462 were HIV-1 uninfected, 4 were HIV-1 infected, and 21 were of indeterminate HIV-1 infection status. Of the 6 infants who died during the first 6 months of life, all were of indeterminate HIV-1 infection status. Three of those 6 deaths occurred during the neonatal period (2 attributable to necrotizing enterocolitis and 1 attributable to perinatal asphyxia), and 3 deaths occurred after the first 1 month of life (attributable to sepsis, pneumonia, and bronchoaspiration of the gastric contents). Therefore, the study population included 462 infants who were known to be HIV-1 uninfected and who completed the 6-month study visit.

Characteristics of the Study Population and Occurrence of Infections

Characteristics of the study population are shown in Table 1. None of the infants was breastfed. All except 1 infant received zidovudine prophylaxis (median duration: 6.0 weeks; 5th percentile: 6.0 weeks; 95th percentile: 8.0 weeks). The remaining infant received nevirapine prophylaxis. Of 462 infants with 11 644 child-weeks of observation, 283 infants (61%) had ≥1 infection; 121 (43%) of 283 infants had ≥1 infection with onset during the neonatal period and 218 (77%) of 283 children had ≥1 infection with onset during the postneonatal period. The 283 infants experienced 526 infections (mean: 1.9 events per infant). Few infants had >2 infections (1.3% during the early neonatal period, 4.7% during the late neonatal period, and 14.5% during the postneonatal period). The overall incidence rate of infections was 4.5 infections per 100 child-weeks of observation (95% CI: 4.1–4.7 infections per 100 child-weeks of observation).

Types of Infections

Overall, the most common infections were skin or mucous membrane infections and lower respiratory tract infections.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No.</th>
<th>Infection (n = 119 (25.7%)), n (%)</th>
<th>Univariate OR (95% CI)</th>
<th>P&lt;sub&gt;a&lt;/sub&gt;</th>
<th>Infection (n = 217 (46.4%)), n (%)</th>
<th>Univariate OR (95% CI)</th>
<th>P&lt;sub&gt;a&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal country of residence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>325</td>
<td>87 (26.8)</td>
<td>1.0</td>
<td></td>
<td>164 (50.5)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>101</td>
<td>22 (21.8)</td>
<td>0.8 (0.4–1.3)</td>
<td></td>
<td>39 (38.6)</td>
<td>0.6 (0.4–1.0)</td>
<td></td>
</tr>
<tr>
<td>Bahamas</td>
<td>20</td>
<td>5 (25.0)</td>
<td>0.9 (0.3–2.6)</td>
<td>.7</td>
<td>12 (60.0)</td>
<td>1.8 (0.6–3.7)</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>16</td>
<td>5 (31.3)</td>
<td>1.2 (0.4–3.7)</td>
<td></td>
<td>2 (12.5)</td>
<td>0.1 (0.03–0.6)</td>
<td>.003</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal education, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥13</td>
<td>18</td>
<td>5 (27.8)</td>
<td>1.0</td>
<td></td>
<td>3 (16.7)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>7–12</td>
<td>265</td>
<td>69 (26.0)</td>
<td>0.9 (0.3–2.7)</td>
<td></td>
<td>123 (46.4)</td>
<td>4.3 (1.2–15.3)</td>
<td></td>
</tr>
<tr>
<td>0–6</td>
<td>179</td>
<td>45 (25.1)</td>
<td>0.9 (0.3–2.6)</td>
<td>.9</td>
<td>91 (50.8)</td>
<td>5.2 (1.4–18.5)</td>
<td>.02</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gainfully employed outside the home</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>109</td>
<td>29 (26.6)</td>
<td>1.0</td>
<td></td>
<td>57 (52.3)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>353</td>
<td>90 (25.5)</td>
<td>0.9 (0.6–1.5)</td>
<td>.8</td>
<td>160 (45.3)</td>
<td>0.8 (0.5–1.2)</td>
<td>.2</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No of persons living in household</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–3</td>
<td>220</td>
<td>51 (23.2)</td>
<td>1.0</td>
<td></td>
<td>100 (45.5)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>≥4</td>
<td>242</td>
<td>68 (28.1)</td>
<td>1.3 (0.9–2.0)</td>
<td>.2</td>
<td>117 (48.4)</td>
<td>1.1 (0.8–1.6)</td>
<td>.6</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tobacco use during pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>359</td>
<td>83 (23.1)</td>
<td>1.0</td>
<td></td>
<td>171 (47.6)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>103</td>
<td>36 (35.0)</td>
<td>1.8 (1.1–2.9)</td>
<td>.02</td>
<td>46 (44.7)</td>
<td>0.9 (0.6–1.4)</td>
<td>.7</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol use during pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>414</td>
<td>101 (24.4)</td>
<td>1.0</td>
<td></td>
<td>192 (46.4)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>48</td>
<td>18 (37.5)</td>
<td>1.9 (1.0–3.5)</td>
<td>.06</td>
<td>25 (52.1)</td>
<td>1.3 (0.7–2.3)</td>
<td>.5</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marijuana use during pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>452</td>
<td>116 (25.7)</td>
<td>1.0</td>
<td></td>
<td>211 (46.7)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>10</td>
<td>3 (30.0)</td>
<td>1.2 (0.3–4.9)</td>
<td>.7</td>
<td>6 (60.0)</td>
<td>1.7 (0.5–6.2)</td>
<td>.5</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cocaine use during pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>454</td>
<td>116 (25.5)</td>
<td>1.0</td>
<td></td>
<td>212 (46.7)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>8</td>
<td>3 (37.5)</td>
<td>1.8 (0.4–7.4)</td>
<td>.4</td>
<td>5 (62.5)</td>
<td>1.9 (0.5–8.1)</td>
<td>.5</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reason for antiretroviral receipt during pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prophylaxis</td>
<td>261</td>
<td>56 (21.5)</td>
<td>1.0</td>
<td></td>
<td>110 (42.2)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>195</td>
<td>61 (31.3)</td>
<td>1.7 (1.1–2.5)</td>
<td>.02</td>
<td>105 (53.9)</td>
<td>1.6 (1.1–2.3)</td>
<td>.01</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal HIV-1 classification at hospital discharge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>383</td>
<td>85 (22.2)</td>
<td>1.0</td>
<td></td>
<td>177 (46.2)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>38</td>
<td>21 (53.3)</td>
<td>4.3 (2.2–8.6)</td>
<td></td>
<td>23 (60.5)</td>
<td>1.8 (0.9–3.5)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>41</td>
<td>13 (31.7)</td>
<td>1.6 (0.8–3.3)</td>
<td>.00001</td>
<td>17 (41.5)</td>
<td>0.8 (0.4–1.6)</td>
<td>.2</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; cell count at hospital discharge, cells per mm&lt;sup&gt;3&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥500</td>
<td>175</td>
<td>40 (22.9)</td>
<td>1.0</td>
<td></td>
<td>72 (41.1)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>200 to &lt;500</td>
<td>207</td>
<td>61 (29.5)</td>
<td>1.4 (0.9–2.2)</td>
<td></td>
<td>113 (54.6)</td>
<td>1.7 (1.3–2.6)</td>
<td></td>
</tr>
<tr>
<td>&lt;200</td>
<td>42</td>
<td>10 (23.8)</td>
<td>1.1 (0.5–2.3)</td>
<td>.3</td>
<td>23 (54.8)</td>
<td>1.7 (0.9–3.4)</td>
<td>.02</td>
</tr>
<tr>
<td>Missing</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasma HIV-1 RNA concentration at hospital discharge, copies per mL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000 to &lt;10000</td>
<td>335</td>
<td>90 (26.9)</td>
<td>1.0</td>
<td></td>
<td>164 (49.0)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>≥10000</td>
<td>37</td>
<td>10 (27.0)</td>
<td>1.0 (0.5–2.2)</td>
<td>1.0</td>
<td>20 (54.1)</td>
<td>1.2 (0.6–2.4)</td>
<td>.7</td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode of delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaginal</td>
<td>188</td>
<td>41 (21.8)</td>
<td>1.0</td>
<td></td>
<td>81 (43.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cesarean section type Ib</td>
<td>181</td>
<td>51 (28.2)</td>
<td>1.4 (0.9–2.3)</td>
<td></td>
<td>94 (51.9)</td>
<td>1.4 (0.9–2.2)</td>
<td></td>
</tr>
<tr>
<td>Cesarean section type II</td>
<td>90</td>
<td>26 (28.9)</td>
<td>1.5 (0.8–2.6)</td>
<td>.3</td>
<td>41 (45.5)</td>
<td>1.1 (0.7–1.8)</td>
<td>.2</td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
infections (Table 2). Infections of the respiratory tract (upper and lower) occurred among 150 infants (32.5%) (data not shown). The most infrequent types of infections were systemic infections, congenital infections, and infections of the central nervous system, gastrointestinal tract, and renal/urinary tract (Table 2). Early neonatal sepsis was diagnosed for 12 infants (26.0 cases per 1000 infants; 95% CI: 14.1–46.2 cases per 1000 infants).

Hospitalizations
Table 3 summarizes hospitalizations and delayed hospital discharge (>3 days after birth) attributable to infection. Overall, 81 (17.5%) of 462 infants were hospitalized at least once with an infection. The proportions of infants with hospitalization or delayed hospital discharge attributable to an infection were as follows: early neonatal, 25 infants (5.4%); late neonatal, 14 infants (3.0%); postneonatal, 44 infants (9.5%). None of the infants with upper respiratory tract infections were hospitalized. In contrast, infants with lower respiratory tract infections were hospitalized frequently (40.7%). Overall, of the 108 episodes of lower respiratory tract infections, 20 (18.5%) were attributable to pneumonia (12 hospitalized; 60.0%) and 88 (81.5%) were attributable to bronchiolitis (33 hospitalized; 37.5%). In contrast, there were only 4 hospitalizations (11.8%) among 34 episodes of acute gastroenteritis.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No.</th>
<th>Infection (n = 119 [25.7%]), n (%)</th>
<th>Univariate OR (95% CI)</th>
<th>P&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Infection (n = 217 [46.4%]), n (%)</th>
<th>Univariate OR (95% CI)</th>
<th>P&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal receipt of intrapartum antibiotic treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>239</td>
<td>53 (22.2)</td>
<td>1.0</td>
<td></td>
<td>97 (40.6)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>205</td>
<td>57 (27.8)</td>
<td>1.4 (0.9–2.1)</td>
<td>.2</td>
<td>113 (55.1)</td>
<td>1.8 (1.2–2.6)</td>
<td>.002</td>
</tr>
<tr>
<td>Missing</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>225</td>
<td>49 (21.8)</td>
<td>1.0</td>
<td></td>
<td>107 (47.5)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>234</td>
<td>69 (29.5)</td>
<td>1.5 (1.0–2.3)</td>
<td>.07</td>
<td>109 (46.6)</td>
<td>1.0 (0.7–1.4)</td>
<td>.9</td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gestational age, wk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥37</td>
<td>410</td>
<td>99 (24.2)</td>
<td>1.0</td>
<td></td>
<td>188 (45.8)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>&lt;37</td>
<td>49</td>
<td>19 (38.8)</td>
<td>2.0 (1.1–3.7)</td>
<td>.04</td>
<td>28 (57.1)</td>
<td>1.6 (0.9–2.9)</td>
<td>.2</td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2500</td>
<td>395</td>
<td>92 (23.3)</td>
<td>1.0</td>
<td></td>
<td>184 (46.6)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>&lt;2500</td>
<td>64</td>
<td>26 (40.6)</td>
<td>2.3 (1.3–3.9)</td>
<td>.005</td>
<td>32 (50.0)</td>
<td>1.1 (0.7–1.9)</td>
<td>.7</td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight-for-age percentile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥5th</td>
<td>340</td>
<td>78 (22.9)</td>
<td>1.0</td>
<td></td>
<td>154 (45.3)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>&lt;5th</td>
<td>115</td>
<td>40 (34.8)</td>
<td>1.8 (1.1–2.8)</td>
<td>.01</td>
<td>62 (53.9)</td>
<td>1.4 (0.9–2.2)</td>
<td>.1</td>
</tr>
<tr>
<td>Missing</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant hemoglobin level at hospital discharge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>281</td>
<td>63 (22.4)</td>
<td>1.0</td>
<td></td>
<td>132 (47.0)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Abnormal</td>
<td>104</td>
<td>38 (36.5)</td>
<td>2.0 (1.2–3.2)</td>
<td>.006</td>
<td>59 (56.7)</td>
<td>1.5 (0.9–2.3)</td>
<td>.1</td>
</tr>
<tr>
<td>Missing</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant absolute lymphocyte count at hospital discharge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abnormal</td>
<td>1</td>
<td>0</td>
<td>1 (1.0)</td>
<td></td>
<td>1 (100.0)</td>
<td>1 (1.0)</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>382</td>
<td>26.7</td>
<td>1.0</td>
<td></td>
<td>192 (50.3)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant absolute neutrophil count at hospital discharge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>347</td>
<td>94 (27.1)</td>
<td>1.0</td>
<td></td>
<td>174 (50.1)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Abnormal</td>
<td>13</td>
<td>3 (23.1)</td>
<td>0.8 (0.2–3.0)</td>
<td>.10</td>
<td>8 (61.5)</td>
<td>1.6 (0.5–5.0)</td>
<td>.6</td>
</tr>
<tr>
<td>Missing</td>
<td>102</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congenital anomalies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>428</td>
<td>109 (25.5)</td>
<td>1.0</td>
<td></td>
<td>198 (46.3)</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>31</td>
<td>9 (29.0)</td>
<td>0.8 (0.4–1.9)</td>
<td>.7</td>
<td>18 (58.1)</td>
<td>0.6 (0.3–1.3)</td>
<td>.3</td>
</tr>
<tr>
<td>Missing</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> P values were obtained by using the Fisher-Freeman-Halton exact test.

<sup>b</sup> Cesarean section was performed before labor and before ruptured membranes.

<sup>c</sup> Cesarean section was performed after labor and/or after ruptured membranes.
TABLE 2  Numbers and Types of Infections, Age at Onset of Infection, and Incidence Rate of Infections for Infants With ≥1 Infection

<table>
<thead>
<tr>
<th>Diagnostic Category</th>
<th>No. (%) of Infants</th>
<th>No. (%) of Infections</th>
<th>Incidence Rate (95% CI), Infections per 100 Child-Weeks at Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0–6 d×</td>
<td>7–27 d</td>
<td>28–60 d</td>
</tr>
<tr>
<td>Skin, mucous membraneab</td>
<td>12</td>
<td>61</td>
<td>52</td>
</tr>
<tr>
<td>Lower respiratory tractc</td>
<td>2</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>Upper respiratory tractc</td>
<td>0</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>Systemic</td>
<td>12</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Congenital</td>
<td>11</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Central nervous system</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Renal and urinary system</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

a Age at onset.
b A total of 155 infants had 232 infections involving the skin or mucous membranes; 98 infants had 1 infection; 44 infants had 2 infections; 9 infants had 3 infections; 2 infants had 4 infections; 1 infant had 5 infections; and 1 infant had 6 infections. The types of infections were as follows: oral Candida infection, 167; cutaneous candidiasis, 21; impetigo, 16; scabies, 12; dermatitis, 8; cellulitis, 2; lymphadenitis, 1; vulvovaginitis, 1; tinea, 1; gingivitis, 1; gingivostomatitis, 1; pityriasis, 1.

c Ninety-five infants had 108 infections involving the lower respiratory tract; 84 infants had 1 infection; 9 infants had 2 infections; and 2 infants had 3 infections. The types of infections were as follows: bronchiolitis, 72; bronchitis, 16; pneumonia, 20.

d Seventy-three infants had 91 infections involving the upper respiratory tract; 58 infants had 1 infection; 12 infants had 2 infections; and 3 infants had 3 infections. The types of infections were as follow: upper respiratory tract infection, not otherwise specified, 52; otitis media, 21; tonsillitis/pharyngitis, 14; viral rhinitis, 1; otitis externa, 3.

e Forty infants had 67 infections involving the gastrointestinal tract; 26 infants had 1 infection; 9 infants had 2 infections; and 4 infants had 3 infections. The types of infections were as follows: diarrhea, 22; acute gastroenteritis, 11; hepatitis C, 1.

Analyses of Infections in the Neonatal Period (Onset at <28 Days of Age)

Variables that were associated in univariate analyses (P ≤ .2) with an increased risk of infections beginning in the neonatal period and were included initially in the multivariate modeling were number of persons living in the household, maternal tobacco use during pregnancy, maternal alcohol use during pregnancy, infant anemia at birth, and number of persons living in the household (crowding) at birth.

Analyses of Infections in the Postneonatal Period (Onset at ≥28 Days of Age)

Variables that were associated in univariate analyses (P ≤ .2) with an increased risk of infections beginning in the postneonatal period and were included initially in the multivariate modeling were maternal country of residence, maternal education, maternal employment, reason for receipt of antiretroviral drugs during pregnancy, infant anemia at birth, and number of persons living in the household (crowding) at birth.

TABLE 3  Number of Hospitalizations or Delayed (>3 Days) Hospital Discharges After Birth Attributable to Infections, According to Diagnostic Category and Age

<table>
<thead>
<tr>
<th>Diagnostic Category</th>
<th>Hospitalizations or Delayed Hospital Discharges, n/N (%) [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early Neonatal (0–6 d)</td>
</tr>
<tr>
<td></td>
<td>Late Neonatal (7–28 d)</td>
</tr>
<tr>
<td></td>
<td>Postneonatal (&gt;28 d)</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Skin and mucous membrane a</td>
<td>0/12 (0.0)</td>
</tr>
<tr>
<td></td>
<td>0/76 (0.0)</td>
</tr>
<tr>
<td></td>
<td>1/140 (0.7)</td>
</tr>
<tr>
<td></td>
<td>0/4 (0.0)</td>
</tr>
<tr>
<td></td>
<td>1/232 (0.4)</td>
</tr>
<tr>
<td>Lower respiratory tract a</td>
<td>2/2 (100.0)</td>
</tr>
<tr>
<td></td>
<td>8/14 (57.1)</td>
</tr>
<tr>
<td></td>
<td>34/90 (37.8)</td>
</tr>
<tr>
<td></td>
<td>0/2 (0.0)</td>
</tr>
<tr>
<td></td>
<td>44/108 (40.7)</td>
</tr>
<tr>
<td>Upper respiratory tract a</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>0/5 (0.0)</td>
</tr>
<tr>
<td></td>
<td>0/80 (0.0)</td>
</tr>
<tr>
<td></td>
<td>0/6 (0.0)</td>
</tr>
<tr>
<td></td>
<td>0/91 (0.0)</td>
</tr>
<tr>
<td>Systemic</td>
<td>11/12 (91.7)</td>
</tr>
<tr>
<td></td>
<td>3/42 (92.9)</td>
</tr>
<tr>
<td></td>
<td>2/14 (13.3)</td>
</tr>
<tr>
<td></td>
<td>0/2 (0.0)</td>
</tr>
<tr>
<td></td>
<td>16/34 (47.1)</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>1/1 (100.0)</td>
</tr>
<tr>
<td></td>
<td>0/1 (0.0)</td>
</tr>
<tr>
<td></td>
<td>3/30 (10.0)</td>
</tr>
<tr>
<td></td>
<td>0/2 (0.0)</td>
</tr>
<tr>
<td></td>
<td>4/34 (11.8)</td>
</tr>
<tr>
<td>Congenital</td>
<td>11/11 (100.0)</td>
</tr>
<tr>
<td></td>
<td>1/3 (33.3)</td>
</tr>
<tr>
<td></td>
<td>1/2 (50.0)</td>
</tr>
<tr>
<td></td>
<td>0/0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>13/16 (81.2)</td>
</tr>
<tr>
<td>Central nervous system a</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>2/3 (66.7)</td>
</tr>
<tr>
<td></td>
<td>1/1 (100.0)</td>
</tr>
<tr>
<td></td>
<td>0/0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>3/4 (75.0)</td>
</tr>
<tr>
<td>Renal and urinary system</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>2/3 (66.7)</td>
</tr>
</tbody>
</table>

PEDIATRICS Volume 119, Number 3, March 2007 e699
receipt of intrapartum antibiotic treatment, infant gestational age, infant weight-for-age percentile at birth, and infant hemoglobin level at hospital discharge (Table 1). Estimated ORs and CIs, both unadjusted and adjusted, for the variables included in the final model for postneonatal infections are shown in Table 5. After adjustment, 3 variables remained in the model, namely, lower maternal CD4+ cell count at hospital discharge, maternal receipt of intrapartum antibiotic treatment, and maternal country of residence.

**DISCUSSION**

In this large cohort of HIV-1–exposed but uninfected infants in Latin America and the Caribbean region, ~60% of infants experienced infectious disease morbidity during the first 6 months of life. Lower respiratory tract and systemic infections frequently were severe, because they were the cause of hospitalization for many infants. More-advanced maternal HIV-1 disease, maternal tobacco use during pregnancy, infant anemia at birth, and number of persons living in the household were associated with neonatal infections, whereas lower maternal CD4+ cell counts soon after delivery, maternal receipt of intrapartum antibiotics, and country of residence were associated with postneonatal infections.

Although there have been studies of the natural history of pediatric HIV-1 infections and associated morbidity, only a few prospective studies in North America, Europe, and Africa have assessed morbidity among HIV-1–exposed but uninfected infants, compared with HIV-1–infected infants. However, none of those studies focused specifically on infectious disease morbidity, which potentially can be preventable.

We found a much higher incidence of neonatal infections (119 of 468 infants [26%] experienced ≥1 infection) than did a large North American study, in which 4.2% of infants <30 days of age had ≥1 infection. The higher incidence of presumed early neonatal sepsis (26.0 cases per 1000 infants), compared with that found for the North American population (1.0–8.1 cases per 1000 infants), is also concerning.

Our observed rate of presumed oral candidiasis (which represented approximately two thirds of the skin and mucous membranes infections) was comparable to

### Table 4

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>OR (95% CI)</th>
<th>Adjusted OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal clinical disease stage at hospital discharge (CDC clinical classification)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>B</td>
<td>4.3 (2.2–8.6)</td>
<td>5.4 (2.6–11.3)</td>
</tr>
<tr>
<td>C</td>
<td>1.6 (0.8–3.3)</td>
<td>2.1 (0.9–4.5)</td>
</tr>
<tr>
<td>Maternal tobacco use during pregnancy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Yes</td>
<td>1.8 (1.1–2.9)</td>
<td>2.2 (1.3–3.8)</td>
</tr>
<tr>
<td>Infant hemoglobin level at hospital discharge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Abnormal</td>
<td>2.0 (1.2–3.2)</td>
<td>2.3 (1.4–3.8)</td>
</tr>
<tr>
<td>No. of persons living in household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–3</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>≥4</td>
<td>1.3 (0.9–2.0)</td>
<td>1.5 (0.9–2.4)</td>
</tr>
</tbody>
</table>

CDC indicates Centers for Disease Control and Prevention.

*ORs and CIs were adjusted for all 4 variables.

### Table 5

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>OR (95% CI)</th>
<th>Adjusted OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal CD4+ count at hospital discharge after delivery, cells per mm3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥500</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>200 to &lt;499</td>
<td>1.7 (1.1–2.6)</td>
<td>1.9 (1.2–2.9)</td>
</tr>
<tr>
<td>&lt;200</td>
<td>1.7 (0.9–3.4)</td>
<td>1.7 (0.8–3.6)</td>
</tr>
<tr>
<td>Maternal receipt of intrapartum antibiotic treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Yes</td>
<td>1.8 (1.2–2.6)</td>
<td>1.7 (1.2–2.5)</td>
</tr>
<tr>
<td>Country</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Argentina</td>
<td>0.6 (0.4–1.0)</td>
<td>0.5 (0.3–0.8)</td>
</tr>
<tr>
<td>Bahamas</td>
<td>1.8 (0.6–3.7)</td>
<td>1.3 (0.6–3.9)</td>
</tr>
<tr>
<td>Mexico</td>
<td>0.1 (0.03–0.6)</td>
<td>0.1 (0.02–0.4)</td>
</tr>
</tbody>
</table>

*ORs and CIs were adjusted for all 3 variables.
that observed among HIV-1–uninfected African children <18 months of age. In addition, although the incidence of diarrhea and gastroenteritis (7.1%) among HIV-1–uninfected infants in our study, who did not have the benefits of breastfeeding, was similar to the estimates for Brazilian infants born to HIV-1–uninfected mothers, it was much lower than that (35%) found for infants <6 months of age in Brazilian urban slums and that (30%) for HIV-1–uninfected infants from South Africa. All HIV-1–exposed infants enrolled in the NISDI Perinatal Study had replacement feeding ensured and special attention regarding hygienic measures for formula preparation, which might have affected positively the incidence and severity of diarrhea but not oral candidiasis among these infants.

Our findings of any respiratory infection for 150 of the studied infants (32.5%) was higher than that in a Brazilian community-based study (20%). Although there is the possibility of recall bias with respect to upper respiratory tract infections, this is much less likely for lower respiratory tract infections. Recent data from community-based studies regarding lower respiratory tract infection rates in normal, HIV-1–unexposed infants are not available. Existing estimates based on large sample sizes for infants <6 months of age in the general population vary widely, from 0.08 to 0.57 cases per 100 child-weeks at risk in the United States to 426.4 to 1591 cases per 100 child-weeks at risk in developing countries. Our findings (0.7–1.1 cases per 100 child-weeks) are closer to those of developed countries.

Ideally, a comparison group of infants born to HIV-1–uninfected mothers from a similar sociodemographic background would be included in studies to evaluate whether HIV-1–uninfected infants born to HIV-1–infected mothers are more susceptible to lower respiratory tract infections than are infants without HIV-1–infected mothers. Other authors evaluating HIV-1–uninfected infants have not included a control group of normal infants. HIV-1–uninfected infants in our study had a rate of lower respiratory tract infections similar to that of South African, HIV-1–infected infants <18 months of age (15.3%). As in a North American study, approximately two thirds of the lower respiratory tract infections observed in our study were diagnosed as acute bronchiolitis and the rest as pneumonia. Considering the young age, it is likely that most of these infections were caused by viruses, mainly respiratory syncytial virus (as confirmed for 1 infant), one of the most common and troublesome viruses of infancy associated with acute bronchiolitis.

Infants with lower respiratory tract infections had a very high (41%) hospitalization rate. In a Brazilian cohort of 5304 infants, only 2% were hospitalized because of acute bronchiolitis. In general, high hospitalization rates have been reported only for infants with underlying conditions, such as prematurity, congenital heart disease, or bronchopulmonary dysplasia. The highest hospitalization rate for infants with bronchopulmonary dysplasia in a developed country was 39%. It is possible that the very high hospitalization rate observed in our study of term infants without underlying conditions could be explained by the hospital admission practices for HIV-1–uninfected infants at the research study sites. However, because higher hospitalization rates may correspond to more-severe disease, with the possibility of long-term sequelae, it is important to evaluate further whether HIV-1–exposed but uninfected infants are at high risk for bronchiolitis, as well as to determine what factors, besides lack of breastfeeding, are related to infection and disease severity.

In terms of risk factors for infections among these HIV-1–exposed but uninfected infants, the association of more-advanced maternal HIV-1 clinical disease stage with neonatal infections leads to the hypothesis that high rates of maternal genital colonization with pathogens or subclinical chorioamnionitis among HIV-1–infected mothers because of immunologic deterioration may play a role in early-onset sepsis. In addition, HIV-1–infected women with advanced HIV-1 disease may have lower antibody titers for common pathogens and/or a lower transfer of these protective antibodies across the placenta. Similarly, the association of low maternal CD4+ cells and postneonatal infections may reflect lower protective titers of passively acquired antibodies and exposure to mothers highly colonized with pathogens because of immunodeficiency. Indeed, a higher risk of morbidity during the first few months of life among African HIV-1–exposed but uninfected infants born to HIV-1–infected mothers with advanced disease has been described.

Infant anemia has been associated with infectious gastrointestinal and respiratory morbidity in young infants from a developing country, even after controlling for environmental and socioeconomic factors. The identification of anemia at birth as a risk factor for neonatal infections reinforces the need for close monitoring of hemoglobin values among infants of HIV-1–infected mothers, especially those whose mothers received antiretroviral drugs during pregnancy.

It is well known that maternal smoking during pregnancy affects the fetus in a number of ways that may result in chronic hypoxia, metabolic changes, and low birth weight. Aside from adverse effects on pulmonary function in the neonatal period, there have been no reports of maternal smoking during pregnancy and neonatal infections (as has been shown for alcohol use during pregnancy).

It is well recognized that postneonatal infections are associated with the socioeconomic status of the mother (and this may vary substantially from country to country). We do not know the mechanism for the association
between maternal receipt of intrapartum antibiotic treatment and postneonatal infections.

The strengths of this analysis include the use of prospectively collected data beginning during pregnancy. To avoid underestimation of the incidence of infections, both presumed and confirmed infections were analyzed. We decreased the likelihood of misclassification bias by avoiding classification of any early-onset respiratory condition without systemic signs or typical radiologic findings as newborn pneumonia, although this approach could have underestimated the rates of pneumonia and/or early-onset sepsis.

A limitation of our study, along with most or all other studies in this area, is that definitive comparisons with infants born to HIV-1–uninfected mothers from the same population could not be made because we did not have a control group (which ideally would be composed of nonbreastfed infants). Selection bias is possible, because of enrollment of mothers who were willing to enroll in a research study and were willing to be monitored, with their infants, for several months. Finally, in our cohort, most systemic neonatal infections were clinical sepsis. Taking into account the fact that clinical sepsis can be also caused by noninfectious perturbations of homeostasis, the rate of bloodstream infections found in this study could be an overestimate.

CONCLUSIONS

We have provided the first characterization of infectious disease morbidity in a large cohort of HIV-1–exposed infants in Latin America and the Caribbean region. Of particular concern is the frequent incidence of early neonatal sepsis and mucocutaneous candidiasis and hospitalizations for acute bronchiolitis. Close monitoring of HIV-1–exposed infants, especially those who are anemic at birth or whose mothers have more-advanced HIV-1 disease or who smoke during pregnancy, remains important. Knowledge of such risk factors for infectious disease morbidity during the neonatal and postneonatal periods may facilitate development of appropriate clinical interventions to decrease the frequency and severity of infectious diseases among these infants.

ACKNOWLEDGMENTS

This work was supported by NICHD contract N01-HD-3-3345.

The NISDI Perinatal Study Group principal investigators, study coordinators, coordinating center representatives, and NICHD staff members include the following: Marcelo H. Losso, Adriana S. Duran, Silvina Ivalo, Alejandro Hakim (Hospital General de Agudos Dra Cecilia Grierson, Buenos Aires, Argentina); Perry Gomez, Percival McNeil, Marva Jervis, Channelle Diggiss, Rosamay Bain (Princess Margaret Hospital, Nassau, Bahamas); Jorge Pinto, Victor Melo, Fabiana Kakehasi (Universidade Federal de Minas Gerais, Belo Horizonte, Brazil); Ricardo de Souza, Jose Mauro Madi (Universidade de Caxias do Sul/Hospital Geral de Caxias do Sul, Caxias do Sul, Brazil); Ricardo de Souza, Rosangela Boff, Ruti Pipi (Universidade de Caxias do Sul/Ambulatório Municipal de Doencas Sexuamente Transmissives/AIDS, Caxias do Sul, Brazil); Ricardo de Souza, Breno Riegel Santos, Rita Lira (Universidade de Caxias do Sul/Hospital Concejai, Porto Alegre, Brazil); Ricardo de Souza, Rosana da Fonseca, Mario Peixoto, Rita Lira (Universidade de Caxias do Sul/Hospital Femina, Porto Alegre, Brazil); Marisa M. Mussi-Pinhata, Geraldo Duarte, Alessandra C. Marcolin (Hospital das Clinicas da Faculade de Medicina de Ribeirao Preto da Universidade de Sao Paulo, Ribeirao Preto, Brazil); Marcos Machado D’Ippolito, Esau Custodio Joao, Jacqueline Menezes, Guilherme Amaral Calvet (Hospital dos Servidores do Estado, Rio de Janeiro, Brazil); Regina Celia de Menezes Succi, Prescilla Chow Lindsey (Federal University of Sao Paulo, Sao Paulo, Brazil); Javier Ortiz Ibarra, Ricardo Figueroa-Damian, Guadalupe Noemi Plazola-Camacho (Instituto Nacional de Perinatologia, Mexico City, Mexico); Data Management and Statistical Center: René Gonin, James Korelitz, Susan Truitt, Roslyn Hennessey, Yolanda Bertucci, Laura Freimanis, D. Robert Harris, Julianne Byrne (Westat); NICHD: Lynne Mofenson, Jack Moye, Jennifer S. Read, Leslie Serchuck, Heather Watts.

REFERENCES

**APPENDIX Infectious Disease Diagnoses Included in Each Category and Protocol Definitions for Each Diagnosis**

<table>
<thead>
<tr>
<th>Category</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congenital</td>
<td>Congenital syphilis, early, presumed: &lt;1 y of age, based on maternal history, infant or maternal serologic findings, and clinical presentation of infant; organism not detected</td>
</tr>
<tr>
<td></td>
<td>Congenital syphilis, late, symptomatic: &gt;1 y of age, seropositive, with clinical evidence of late sequelae of congenital syphilis</td>
</tr>
<tr>
<td></td>
<td>Prophylaxis of congenital syphilis</td>
</tr>
<tr>
<td></td>
<td>Toxoplasmosis: diagnosed on the basis of IgM serologic or histopathologic findings within 1–5 mo of life, with no symptoms or signs</td>
</tr>
<tr>
<td>Central nervous system</td>
<td>Encephalitis, proven: pathogen identified with PCR assays, serologic assays, or biopsy</td>
</tr>
<tr>
<td></td>
<td>Meningitis, presumed: suspected clinically; abnormal CSF findings consistent with diagnosis; tests negative for specific organism in CSF or blood</td>
</tr>
<tr>
<td></td>
<td>Meningitis, proven: abnormal CSF findings consistent with diagnosis; identification of specific organism in CSF or blood</td>
</tr>
<tr>
<td>Upper respiratory tract</td>
<td>Otitis externa</td>
</tr>
<tr>
<td></td>
<td>Otitis media, acute, clinical: diagnosed through physical examination; specific organism not identified</td>
</tr>
<tr>
<td></td>
<td>Otitis media, with effusion: diagnosed through physical examination/tympanometry</td>
</tr>
<tr>
<td></td>
<td>Tonsillitis/pharyngitis, presumed: unknown cause</td>
</tr>
<tr>
<td></td>
<td>Upper respiratory infection</td>
</tr>
<tr>
<td></td>
<td>Viral rhinitis</td>
</tr>
<tr>
<td>Lower respiratory tract</td>
<td>Pneumonia, suspected: fever, tachypnea, cough, and consistent physical findings but chest radiograph not obtained or not available; specific assays negative or not performed</td>
</tr>
<tr>
<td></td>
<td>Pneumonia, presumed: clinical findings and chest radiographic results temporally consistent with diagnosis (within 7 d); specific assays negative or not performed</td>
</tr>
<tr>
<td></td>
<td>Bronchiolitis, presumed: specific cause unknown</td>
</tr>
<tr>
<td></td>
<td>Bronchiolitis, proven: cause proven</td>
</tr>
<tr>
<td></td>
<td>Pneumocystis jirovecii pneumonia: clinical diagnosis: cough, dyspnea, tachypnea, and diffuse bilateral interstitial infiltrates on chest radiograph; laboratory evidence of hypoxia; no evidence of bacterial or viral pneumonia</td>
</tr>
<tr>
<td></td>
<td>Bronchitis, presumed: cause not determined</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>Acute gastroenteritis, presumed: cause not identified</td>
</tr>
<tr>
<td></td>
<td>Diarrhea, presumed: clinical or epidemiologic diagnosis only</td>
</tr>
<tr>
<td>Renal and urinary system</td>
<td>Urinary tract infection, proven: diagnosed through suprapubic tap or catheterization of urine (for children &lt;5 y of age) or clean-catch specimen (for patients &gt;5 y of age) positive for specific organism (excluding pyelonephritis)</td>
</tr>
<tr>
<td>Skin and mucous membranes</td>
<td>Conjunctivitis, proven: cause identified</td>
</tr>
<tr>
<td></td>
<td>Conjunctivitis, presumed: clinical diagnosis</td>
</tr>
<tr>
<td></td>
<td>Dermatitis, other</td>
</tr>
<tr>
<td></td>
<td>Vulvovaginitis, presumed</td>
</tr>
<tr>
<td></td>
<td>Candidiasis, vagina, presumed</td>
</tr>
<tr>
<td></td>
<td>Candidiasis, oral, proven: diagnosed on the basis of gross appearance of white patches or plaques on erythematous base and microscopic appearance of yeast in uncultured specimen scraped from oral mucosa or positive culture</td>
</tr>
<tr>
<td></td>
<td>Cellulitis, presumed: suspected clinically; no organism identified</td>
</tr>
<tr>
<td></td>
<td>Impetigo, presumed: test for specific organism negative or not performed</td>
</tr>
<tr>
<td></td>
<td>Lymphadenitis, presumed: inflammation and swelling of lymph nodes with tenderness; pathogen not identified</td>
</tr>
<tr>
<td></td>
<td>Candida, oral thrush, presumed: diagnosed on the basis of gross appearance of white patches or plaques on erythematous base and response to specific therapy</td>
</tr>
<tr>
<td>Systemic</td>
<td>Sepsis, proven (excluding indwelling catheter culture); identification of pathogen in blood with positive culture</td>
</tr>
<tr>
<td></td>
<td>Sepsis, inflammatory response syndrome: clinical impression of systemic response to bloodstream or loculated infection without documentation of source of infection</td>
</tr>
<tr>
<td></td>
<td>VZV, uncomplicated primary disease (chickenpox); uncomplicated</td>
</tr>
<tr>
<td></td>
<td>HHV-6 infection: suspected infection because of clinical diagnosis of roseola; virus detection studies not performed or negative</td>
</tr>
<tr>
<td></td>
<td>Bacteremia</td>
</tr>
<tr>
<td></td>
<td>Sepsis, line, proven: focus of infection identified as indwelling catheter; identified with positive blood culture or other specific pathogen identification in blood</td>
</tr>
<tr>
<td></td>
<td>Trypanosomiasis</td>
</tr>
<tr>
<td></td>
<td>Rash, nonspecific</td>
</tr>
</tbody>
</table>

CMV indicates cytomegalovirus; PCR, polymerase chain reaction; CSF, cerebrospinal fluid; VZV, varicella zoster virus; HHV, human herpes virus.
Once-Daily Highly Active Antiretroviral Therapy for HIV-Infected Children: Safety and Efficacy of an Efavirenz-Containing Regimen

Henriëtte J. Scherpbier, MD, PhD, Vincent Bekker, MD, PhD, Dasja Pajkrt, MD, PhD, Suzanne Jurriaans, PhD, Joep M. A. Lange, MD, PhD, Taco W. Kuijpers, MD, PhD

Objective. To improve adherence and virologic suppression, we assessed the feasibility and effectiveness of a once-daily regimen of efavirenz with 3 nucleoside reverse transcriptase inhibitors as first-line or second-line highly active antiretroviral therapy in a cohort of HIV-1–infected children.

Methods. HIV-1–infected children naïve to efavirenz were treated with a combination of efavirenz, abacavir, didanosine, and lamivudine in an observational, prospective, single-center study. Virologic failure-free survival was assessed with Kaplan-Meier analysis. The CD4\(^+\) T-cell increase was estimated by using a generalized linear model incorporating repeated measurements.

Results. Thirty-six children received the study medication for a median of 69 weeks. Virologic failure-free survival rates were 76% and 67% after 48 weeks and 96 weeks, respectively. No significant difference was found in efficacy between first-line and second-line highly active antiretroviral therapy. All children receiving highly active antiretroviral therapy showed a sustained CD4\(^+\) T-cell increase, irrespective of virologic suppression. Growth rates improved with highly active antiretroviral therapy. Study medication administration was stopped for 14 children, mostly because of nonadherence (4 cases) or virologic rebound (5 cases) and because of adverse events (unrelated death and grade 2 liver toxicity) in 2 cases. Lipid abnormalities and abacavir-related hypersensitivity were not observed.

Conclusions. For the first time, once-daily highly active antiretroviral therapy is demonstrated to be a safe, convenient, and potent antiretroviral regimen for HIV-1–infected children.
**METHODS**

**Patients**

Between January 2002 and August 2005, a prospective observational study was performed at the Academic Medical Centre (Amsterdam, Netherlands). HIV-1–infected children were eligible when they were 3 months to 18 years of age and had CD4+ T-cell counts of <1750 cells per μL (for those <1 year of age), <1000 cells per μL (for those between 1 and 2 years of age), <750 cells per μL (for those between 3 and 6 years of age), or <500 cells per μL (for those >6 years of age). Previous exposure to ART regimens was allowed. Exclusion criteria consisted of the presence of mutations associated with resistance to efavirenz or to ≥2 of the NRTI study drugs used at the commencement of the once-daily treatment regimen, pregnancy, and HLA typing results unfavorable with respect to abacavir use. All children in our cohort were naïve to all NNRTIs before enrollment. No restrictions were made with respect to ethnicity, gender, route of HIV acquisition, or disease stage. The medical ethics committee of our institute approved the protocol. Parents or caregivers provided written informed consent.

**Medications**

Patients received efavirenz (starting dose according to manufacturer’s instructions) (Table I), abacavir (16 mg/kg; maximum: 600 mg/day), didanosine (200–240 mg/
m²), and lamivudine (4 mg/kg [≤6 weeks] or 8 mg/kg [≥6 weeks]; maximum: 300 mg/day). Dosage adjustments were performed according to the weight of the children. Because efavirenz is not registered for children <3 years of age, pharmacokinetics over 24 hours were tested at day 0, day 14, and 6 weeks for all patients starting efavirenz treatment, followed by consecutive plasma measurements to monitor drug levels for underdosing or overdosing (and to establish a trough level of >1 mg/L, which is considered a target value for virologic success in adults) and children [Kristel Crommentuijn, H.J.S., Alwin Huitema, T.W.K., and Jos Beijnen, manuscript in preparation]).

It was recommended that the children take their regimen with food. When taken as a solution for the optimal treatment of small children in our cohort, didanosine was prepared with acid-binding magnesium hydroxide, according to the instructions of the manufacturer. The didanosine solution was taken either 1 hour before or 1 hour after food was eaten. Otherwise, the extended-release form didanosine-EC was used, without food restriction.

Adherence
The children’s guardians were counseled regarding the importance of treatment adherence. Where appropriate, the children were also counseled accordingly. Members of the treatment team monitored adherence by telephoning the guardians soon after the regimen was started and at each follow-up clinic visit.

Procedures
At each visit, a physical examination was performed, including weight, length, and head circumference measurements. Blood was drawn before and 1 and 2 weeks and 1, 2, and 3 months after initiation of HAART and every 3 to 4 months thereafter. Lymphocyte subsets were analyzed with a FACScan system (Becton Dickinson, San Jose, CA). The pVL was measured with a Versant HIV-1 bDNA 3.0 assay (Bayer, Mijdrecht, Netherlands), with a lower limit of quantification (LLQ) of 50 copies per mL (input: 1 mL of plasma). Virologic failure was defined as 2 consecutive pVL values of ≥50 copies per mL. Patients who never reached a pVL of ≤50 copies per mL were failing at the first measurement after the nadir pVL in the initial decline.

Nucleotide sequence analyses of the HIV-1 protease and RT genes were performed at baseline and after virologic failure. Sequence analyses were performed with the Viroseq HIV-1 genotyping kit (version 2; Abbott Laboratories, Abbott Park, IL). Resistance-conferring mutations were screened as described by the International AIDS Society-USA.

Adverse events were recorded during the study period and were defined as any clinical sign or symptom or meaningful laboratory test abnormality possibly or probably related to the study medication, excluding HIV-related disorders. The National Institute of Allergy and Infectious Diseases toxicity tables were used to grade the severity of pediatric adverse experiences. Parents were asked about the presence of adverse effects at every visit.

Statistical Analyses
The primary outcome was virologic failure-free survival, which was assessed by using Kaplan-Meier analysis. Censoring was applied if the last patient visit or a switch to another regimen occurred before virologic failure. The secondary outcomes were factors associated with virologic failure, changes in CD4⁺ and CD8⁺ T-cell counts over time, changes in growth parameters (weight and height) over time, reported adverse events, and the occurrence of resistance mutations. Age-adjusted CD4⁺ and CD8⁺ T-cell ratios were calculated by dividing the counts by the mean for an age-matched healthy control group. Growth of the children was analyzed by means of the z scores (normal SD) of height and length. These scores were calculated with the use of Growth Analyser 2.0 software (Dutch Growth Foundation, Rotterdam, Netherlands), with Dutch reference values. Age-adjusted CD4⁺ and CD8⁺ T-cell ratios and height and weight z scores were modeled by using a mixed model incorporating repeated measurements. This model handles missing data adequately by estimating the outcome given a specific covariate structure. The estimates of a specific level of the fixed effects were modeled by using the “first-order autoregressive” approach. Differences in these estimates between different levels of the variable were tested for significance by using the t statistic. When subgroups of patients were compared, the differences between groups were evaluated by using Fisher’s exact test for categorical data and the Kruskal-Wallis test for continuous data. All statistical analyses were performed with SPSS for Windows (SPSS, Chicago, IL). A 2-sided P value of <.05 was considered statistically significant.

RESULTS
Patients
All 36 HIV-1–infected children who started a once-daily ART regimen with efavirenz between January 2002 and August 2005 were included in the present analyses. ART-naive and pretreated HIV-1–infected children were included. Twenty-two children (61%) had been receiving HAART for a median 259 weeks before enrollment (interquartile range [IQR]: 104–310 weeks). Of these children, 10 were also pretreated with single/dual NRTI therapy for a median of 134 weeks before the start of HAART. One child used 2 different HAART regimes before enrollment; all others used 1 HAART regimen before enrollment.

Baseline characteristics are shown in Table 2. The median age of the children at baseline was 6.6 years
(IQR: 3.3–10.7 years). One of the children was <1 year of age, 6 were between 1 and 2 years of age, 12 were between 3 and 6 years of age, and 17 were >6 years of age. Children who were naive to ART were younger at baseline than were children who received second-line HAART (median: 3.3 years [IQR: 1.7–9.9 years] vs 8.8 years [IQR: 5.2–11.5 years]; \( P = .04 \)). Thirty-four children (94%) acquired HIV infection perinatally from their HIV-1–infected mothers; 15 children (42%) presented with Centers for Disease Control and Prevention category C–classified AIDS-defining symptoms. The majority of the children were black (African or Surinamese). The children received study medication for a median duration of 69 weeks (IQR: 39–122 weeks).

Virologic Findings
At baseline, the median pVL for the whole group was 3.6 log copies per mL (IQR: 2.4–4.7 log copies per mL). Children who started the once-daily regimen as second-line HAART had a significantly lower pVL than did children who started ART naive to the regimen (median: 2.5 vs 5.4 log copies per mL; \( P < .001 \)). Twelve (55%) of 22 children who started second-line HAART had undetectable levels at the switch of therapy. The virologic failure-free survival rates were 76% and 67% after 48 and 96 weeks, respectively (Fig 1A). Twelve children completed a follow-up period of 96 weeks with study medication. Of the patients who experienced failure of therapy or discontinued study medication at any time during the follow-up period, 6 never had a pVL below the LLQ. Of the remaining 30 children in this observational cohort, 4 showed a rebound of their pVL after having experienced a period of viral suppression below the LLQ after the initiation of study medication. The effect of previous HAART on virologic effectiveness was analyzed with a log-rank test; there was no difference between virologic responders and nonresponders (\( P = .7 \)) (Fig 1B).

Reasons for Treatment Discontinuation
Study medication needed to be discontinued for 14 children (39%) during the follow-up period, for the following reasons: 5 virologic failures with several new mutations, 4 reported cases of nonadherence, 2 cases of aversion to the taste of the medication, 1 pregnancy, 1 serious adverse event (death), and 1 adverse event (grade 2 elevation in liver transaminase levels). The fatality involved a patient who experienced a severe electrolyte disturbance and lactate acidosis because of persistent diarrhea, despite a rapid virologic response to undetectable levels. In the opinion of the treating physicians, this was not attributable to the study drugs. Hypersensitivity to abacavir was not seen.

Resistance Mutations
The \( RT \) gene from HIV-1 in plasma samples was sequenced for all 36 children. The HIV-1 strains in children who experienced failure of the study medication were scrutinized for the occurrence of additional critical mutations in the \( RT \) gene associated with NNRTI resistance, for efavirenz in particular (ie, 100I, 103N, 106A/M, 108I, 181C/I, 190A/S, 225H, or 230L). One HAART-experienced boy had a 181C mutation at the start of the study regimen. His pVL became undetectable with the study medication. For one ART-naive child, a 69N mutation in the \( RT \) gene was found at baseline. Mutations associated with resistance to \( \geq 1 \) NRTI were detected in the group of children who had shown viral blips previously or had

### TABLE 2 Baseline Characteristics of Children Starting Treatment With the Efavirenz-Containing Study Regimen and Comparison Between First-Line and Second-Line HAART

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>First-Line HAART</th>
<th>Second-Line HAART</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>36</td>
<td>14</td>
<td>22</td>
</tr>
<tr>
<td>Female, ( n )</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, median (IQR), y</td>
<td>6.6 (3.3–10.7)</td>
<td>3.3 (1.7–9.9)</td>
<td>8.8 (5.2–11.5)</td>
</tr>
<tr>
<td>CDC-C classification, ( n^a )</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical transmission, ( n )</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual transmission, ( n )</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black, ( n )</td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonblack, ( n )</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD4(^+) T cells, median (IQR), No. per ( \mu L )</td>
<td>730 (400–1050)</td>
<td>460 (170–940)</td>
<td>860 (600–1170)</td>
</tr>
<tr>
<td>CD4(^+) T cells, median (IQR), %</td>
<td>26 (16–37)</td>
<td>14 (6–25)</td>
<td>31 (25–38)</td>
</tr>
<tr>
<td>CD4(^+) T cells, median (IQR), age-adjusted ratio</td>
<td>0.5 (0.3–0.9)</td>
<td>0.3 (0.1–0.5)</td>
<td>0.7 (0.4–0.9)</td>
</tr>
<tr>
<td>CD8(^+) T cells, median (IQR), No. per ( \mu L )</td>
<td>1270 (800–1910)</td>
<td>2060 (790–3480)</td>
<td>1120 (810–1400)</td>
</tr>
<tr>
<td>CD8(^+) T cells, median (IQR), %</td>
<td>44 (34–61)</td>
<td>58 (38–72)</td>
<td>38 (31–49)</td>
</tr>
<tr>
<td>CD8(^+) T cells, median (IQR), age-adjusted ratio</td>
<td>1.5 (1.1–1.8)</td>
<td>1.9 (1.2–2.9)</td>
<td>1.4 (1.0–1.6)</td>
</tr>
<tr>
<td>Total cholesterol level, median (IQR), mmol/L</td>
<td>3.9 (3.4–4.5)</td>
<td>3.4 (3.1–3.6)</td>
<td>4.3 (3.9–4.7)</td>
</tr>
<tr>
<td>Triglyceride level, median (IQR), mmol/L</td>
<td>0.8 (0.6–1.5)</td>
<td>1.2 (0.8–1.7)</td>
<td>0.7 (0.6–1.0)</td>
</tr>
<tr>
<td>HIV-1-RNA, median (IQR), log copies per mL</td>
<td>3.6 (2.4–4.7)</td>
<td>5.4 (4.4–6.0)</td>
<td>2.5 (2.4–3.5)</td>
</tr>
<tr>
<td>Height-for-age z score, median (IQR)</td>
<td>−1.2 (−2.0 to −0.1)</td>
<td>−1.9 (−3.0 to −1.3)</td>
<td>−0.5 (−1.4 to 0.5)</td>
</tr>
<tr>
<td>Weight-for-height z score, median (IQR)</td>
<td>0.6 (−0.5 to 1.2)</td>
<td>0.6 (−0.5 to 1.5)</td>
<td>0.5 (−0.5 to 1.0)</td>
</tr>
</tbody>
</table>

\( ^a \) Clinical categories were as defined by the US Centers for Disease Control and Prevention. 

---

SCHERPBIER et al
experienced complete failure with their first-line, PI-containing, HAART regimen (Table 3). In a survival analysis, there was no significant difference in the time to virologic failure for patients with existing mutations at baseline, compared with children without mutations at baseline ($P = .5$).

### Immunologic Findings

At baseline, the median CD4$^+$ T-cell count for the total study population was 730 cells per $\mu$L (IQR: 400–1050 cells per $\mu$L), the age-adjusted CD4$^+$ T-cell ratio was 0.5 (IQR: 0.3–0.9), and the CD4$^+$ T-cell percentage was 26% (IQR: 16%–37%). The baseline age-adjusted CD4$^+$

![Kaplan-Meier survival analysis of time to virologic failure. Numbers of patients at risk at the beginning and after 1 year are indicated. Censoring was applied if the last patient visit or a switch to a simplified regimen occurred before virologic failure. B, Survival analysis, showing no difference in virologic responders and nonresponders for children receiving first-line HAART (dotted line) and children receiving second-line HAART (solid line) ($P = .7$).]

**FIGURE 1**

A, Kaplan-Meier survival analysis of time to virologic failure. Numbers of patients at risk at the beginning and after 1 year are indicated. Censoring was applied if the last patient visit or a switch to a simplified regimen occurred before virologic failure. B, Survival analysis, showing no difference in virologic responders and nonresponders for children receiving first-line HAART (dotted line) and children receiving second-line HAART (solid line) ($P = .7$).
T-cell ratio and the absolute number and percentage of CD4+ T cells were statistically significantly higher in children who started the regimen as second-line HAART, compared with children who had not received previous ART ($P = .003$, $P = .03$, and $P < .001$, respectively) (Table 2). The median CD8+ T-cell count for the total study population was 1270 cells per μL (IQR: 800–1910 cells per μL), and the age-adjusted CD8+ T-cell ratio was 1.5 (1.1–1.8).

The median age-adjusted CD4+ T-cell ratio demonstrated an increase during the 96 weeks of treatment (Fig 2A). Children who started naive to ART had a more-profound increase, compared with children receiving second-line HAART (Fig 2B). This was attributable to a lower baseline CD4+ T-cell count. The age-adjusted CD8+ T-cell ratios demonstrated slight but nonsignificant decreases in the total study population (Fig 2C) and in both subgroups, based on pretreatment (Fig 2D).

**Lipids**

Although there was a significantly lower total cholesterol level for patients who started naive to ART, compared with patients who started second-line HAART (median: 3.4 vs 4.3 mmol/L; $P < .001$), all children were below the cutoff value of 6.5 mmol/L (upper limit of the
reference range). The same applied to triglyceride levels at baseline (median: 1.1 vs 0.7 mmol/L; \(P = .04\); normal levels: <5.0 mmol/L).

During the treatment with HAART, total cholesterol levels increased. However, for children with second-line HAART, total cholesterol levels remained stable. Children who started naive to ART showed an increase toward the values of the group with second-line HAART within the first weeks. Triglyceride levels did not change over time during treatment with the once-daily regimen.

**Growth and Development**

Growth parameters are shown in Fig 3. The median height-for-age \(z\) score at baseline for the total study population was \(-1.2\), and the median weight-for-height \(z\) score was \(0.6\). Children naive to ART had a significantly lower height-for-age \(z\) score, compared with children receiving second-line HAART (median \(z\) score: \(-1.9\) vs \(-0.5\); \(P = .001\)). The ART-naive group showed a distinct increase in the first 48 weeks, but values did not reach the level of the second-line HAART group (Fig 3B). An increase in weight-for-age \(z\) scores was seen during 96 weeks of treatment, almost to normal values (Fig 3C). Children with second-line HAART showed a different pattern over time, compared with children who started naive to ART (Fig 3D). The children who started naive to ART showed an increase, in contrast to the children receiving second-line HAART, who showed a higher baseline level that remained stable. Weight-for-height \(z\) scores remained stable in both treatment groups (Fig 3, E and F).

**DISCUSSION**

We demonstrated for the first time the virologic effectiveness, tolerability, and safety of a once-daily HAART regimen in an HIV-1–infected pediatric cohort. Virologic failure-free survival rates were 76\% and 67\% after 48 and 96 weeks, respectively, for all children, with equal effectiveness of the study regimen when used as first-line or second-line HAART (77\% and 75\%, respectively, after 48 weeks). The only pediatric report on efavirenz-containing HAART to date involved 17 children with persistently suppressed pVL switching to efavirenz-containing HAART for reasons of convenience and simplification.\(^{13}\) Our treatment regimen contained efavirenz and 3 NRTIs as the backbone and was administered to 36 children. A limited number of studies on once-daily regimens for ART-naive adults have been reported up to 48 weeks, with virologic failure-free survival rates ranging from 50\% to 78\%.\(^{27-32}\) The combination of tenofovir and didanosine as the NRTI backbone seemed the least effective,\(^{30,31}\) despite good adherence as defined with the medication event monitoring system and plasma efavirenz concentration monitoring.\(^{31}\) Our results in chil-
dren demonstrate favorably the strong antiretroviral activity of the chosen once-daily regimen (irrespective of previous treatment experience).

Our efavirenz-containing once-daily regimen was well tolerated. Regarding safety and tolerability, we observed medication-related grade 2 toxicity in only 1 child, consisting of an increase in blood liver enzyme levels. This concurs with other reports on children. \(^{13,17,33}\) The nonfasting lipid levels remained within reference ranges, although total cholesterol levels showed an increase during the first weeks in the ART-naive group, compared with children who started second-line HAART. In the 2NN study, it was shown that the efavirenz-associated increase in total cholesterol levels was mainly attributable to high-density lipoprotein cholesterol. \(^{34}\) One patient with AIDS-related severe cachexia died as a result of persisting electrolyte disturbances despite very successful HIV-1 suppression. For 14 children, the regimen was stopped despite good tolerability and simplification of intake, compared with most of the previous regimens for children. Discontinua-

**FIGURE 3**

A, Height-for-age \(z\) scores during the 96-week follow-up period with HAART. B, Comparison of height-for-age \(z\) scores for children receiving first-line and second-line HAART. C, Weight-for-age \(z\) scores during the 96-week follow-up period with HAART. D, Comparison of weight-for-age \(z\) scores for children receiving first-line and second-line HAART. E, Weight-for-height \(z\) scores during the 96-week follow-up period with HAART. F, Comparison of weight-for-height \(z\) scores for children receiving first-line and second-line HAART. The \(z\) scores were calculated for measurements of height according to age and gender by using the 1997 Dutch reference curves. \(^{39}\) Bars indicate SEM values.
tion was for several reasons but mostly for reasons of virologic rebound attributable to assumed or self-reported nonadherence.

A latent viral reservoir may harbor viruses that are generated at various times throughout the life of perinatally infected children, including wild-type, drug-sensitive viruses transmitted from the mother and any drug-resistant viruses that arise during nonsuppressive HAART therapy. Seventeen of the 22 patients who received second-line HAART showed extensive RT mutations. For one child naïve to ART, a 69N mutation in RT was found at baseline. This mutation is associated with resistance to zidovudine, stavudine, and didanosine. Most probably the virus was acquired from the mother, although the predominant virus population in the mother seemed to contain a 69S mutation in RT. The amino acid difference at this position can be explained by viral evolution, because the baseline sequences of the mother and child were obtained 3.5 years after birth. One HAART-experienced boy had an 181C mutation at the start of the study regimen. Although this mutation is associated with resistance to efavirenz, it has a greater impact on the sensitivity to nevirapine. With the addition of lopinavir/ritonavir to the study regimen, a lasting virologic response was achieved.

Apart from preexisting mutations, the impact of adherence on the effectiveness of HAART must be considered seriously. Of the virologic failures, 2 patients treated previously with HAART, without mutations, reported nonadherence to the regimen by themselves. In most of the other virologic failures, including that of the pregnant girl, efavirenz was repeatedly below the trough level of 1 mg/L even after dose adjustments, which suggests nonadherence. Critical NNRTI-associated resistance mutations were found for 7 of those 8 patients. Similar to the findings of Luzuriaga et al, we did not observe an association between preexisting NRTI (or PI) resistance mutations and success or failure of virologic control (P = .5), which suggests that adherence may indeed be the most important factor for lasting virologic suppression in our cohort.

Increases in CD4+ T-cell counts were observed in both groups, although the increase in cell number was more profound in the group that started naïve to ART, because of the lower baseline counts. As expected, the baseline age-adjusted CD4+ T-cell ratio and absolute and relative CD4+ T-cell counts were statistically significantly higher for children who started the regimen as second-line HAART. No severe clinical infections occurred during the study period in either group, irrespective of virologic failure.

With respect to general growth and development, ART-naïve children showed an increase in height-for-age z scores but did not reach the level of the second-line HAART group at 96 weeks. In our cohort, the children receiving first-line HAART showed normalization of weight-for-age z scores, whereas the second-line HAART group already had almost-normal z scores at the start. Nachman et al described similar findings for 192 children in clinically stable condition, of whom 50% had been pretreated with NRTIs but were naïve to the new HAART regimens. Most of the second-line HAART group in our study had used stavudine, lamivudine, and nelfinavir as first-line HAART, for a median of 259 weeks. Almost 30% had developed clinically evident lipodystrophy.

The follow-up period in the present once-daily study was too short for any effect on fat distribution to be expected. Moreover, the number of ART-naïve children was small, and the total treatment follow-up period was relatively short for judging any alteration in this respect.

CONCLUSIONS

This is the first study of a once-daily, efavirenz-containing, HAART regimen in a pediatric cohort. Our study demonstrated virologic and immunologic effectiveness, even for children who were HAART experienced. We can conclude that the once-daily regimen is a convenient, safe, and robust regimen for children. Most children were ART experienced but demonstrated equal effectiveness of the once-daily regimen, compared with children receiving first-line HAART. Because of the high pVL in young children, a robust regimen is required. A 2-class regimen with 4 drugs containing abacavir was considered more successful. Simplification of the study regimen by stopping one of the NRTIs when the pVL is below the LLQ may be a possibility with similar outcomes and efficacy. A considerable number of patients who stopped the treatment regimen admitted nonadherence or were suspected of nonadherence because of low drug levels in consecutive blood samples, irrespective of the dose adjustments made. Additional improvement of the effectiveness of treatment can be reached in the future, because once-daily regimens allow for various measures to support adherence (such as the medication event monitoring system or directly observed therapy), in larger, well-designed studies on virologic outcomes.

ACKNOWLEDGMENTS

Dr Bekker was supported in part by grant 7006 from the Dutch AIDS Foundation. Bristol-Myers Squibb provided the study medication (didanosine) and Maalox to prepare the didanosine solution for the children in our cohort.

We thank our patients, their parents, and their caregivers for their participation; our HIV nurses, Atie van der Plas and Eugenie le Poole, for their enthusiasm and care of our pediatric cohort; and the home care teams for their support. We thank Kristel Crommentuyn, Rob ter Heine, Alwin Huitema, and Jos Beijnen for their collaboration on the pharmacodynamic population studies.
and for providing the efavirenz levels. We also thank Michael Kangas for critical reading of and comments on the manuscript.

REFERENCES


11. Scherpbier HJ, Bekker V, van Leth F, Jurriaans S, Lange JM, Kuijpers TW. Long-term experience with combination antiretroviral therapy that contains nelfinavir for up to 7 years in a pediatric cohort. Pediatr. 2006;117(3). Available at: www.pediatrics.org/cgi/content/full/117/3/e528


20. Williams J, Ashoe D, Babiker A, Goodall R, Hooker M. FORTE Trial Steering Committee. A virological benefit from an induction/maintenance strategy compared with a standard 3-drug regimen in antiretroviral naive patients: the FORTE Trial. Presented at the 11th Conference on Retroviruses and Opportunistic Infections; February 8–11, 2004; San Francisco, CA


Early-Life Risk Factors for Occurrence of Atopic Dermatitis During the First Year

Mikio Sugiyama, MD, Hirokazu Arakawa, MD, PhD, Kiyoshi Ozawa, MD, PhD, Takahisa Mizuno, MD, PhD, Hiroyuki Mochizuki, MD, PhD, Kenichi Tokuyama, MD, PhD, Akihiro Morikawa, MD, PhD

ABSTRACT

OBJECTIVE. In a prospective birth cohort study, we sought to identify perinatal predictors of the occurrence of atopic dermatitis in the first year of life.

METHODS. Associations of family history, infection during pregnancy, cord blood cytokine concentrations, and skin function parameters with atopic dermatitis were analyzed. Stratum corneum hydration was measured with an impedance meter until 5 days after delivery and again at 1 month.

RESULTS. Complete data were obtained for 213 infants, including 27 diagnosed by a physician as having atopic dermatitis during their first year and 26 diagnosed as having infantile eczema during their first month. The risk of atopic dermatitis during the first year of life was related to maternal atopic dermatitis, lower concentrations of macrophage inflammatory protein-1α in cord blood, and greater skin moisture in the surface and stratum corneum of the forehead and cheek at 1 month of age but not to viral or bacterial infection during pregnancy or breastfeeding. Paternal hay fever was associated negatively with the development of atopic dermatitis. High concentrations of interleukin-5, interleukin-17, and macrophage chemotactic protein-1 and only surface moisture in the cheek were associated with greater risk of infantile eczema in the first month.

CONCLUSIONS. The association of atopic dermatitis in infancy with reduced neonatal macrophage inflammatory protein-1α levels suggests a link with immature immune responses at birth. Stratum corneum barrier disruption in atopic dermatitis may involve impairment of cutaneous adaptation to extrauterine life. The majority of risk factors had different effects on infant eczema and atopic dermatitis, indicating different causes.
A TOPY DERMATITIS (AD), a chronic inflammatory skin disease that usually occurs in the first few years of life,1 has increased dramatically in prevalence in developed countries over the past several decades. AD and other allergic diseases seem to have multifactorial origins, arising from complex interactions between genes and the environment; however, the relative importance of genetic and prenatal environmental factors is not yet clear. AD is one of the earliest manifestations of allergic/ atopic diseases in children. Furthermore, AD is considered a significant risk factor for aeroallergen sensitization at 5 years and a predictor of the subsequent development of asthma.2 Asthma is a worldwide problem, and the disease’s social burden and costs to public and private health care systems are substantial. Therefore, establishing early risk factors for AD may help provide intervention strategies for the primary prevention of asthma.3

Cytokines are considered important regulators of functional maturation in the developing fetal immune system. However, factors that determine the degree of immune competence at birth and during early infancy are not fully understood. Several studies have linked cytokine profiles at birth to subsequent development of allergic disorders.4–7 Tang et al4 found that infants who either exhibited symptoms of atopic disease or had a positive skin test at 1 year of age produced significantly less interferon (IFN)-γ at birth than did infants without atopy. Macaubas et al5 found negative relationships of interleukin (IL)-4, IFN-γ, and tumor necrosis factor (TNF)-α concentrations in cord blood to the risk of asthma, atopy, or both by 6 years of age. These findings suggested that some immune functions, including the capacity to secrete both T helper (Th) type 1 and 2 cytokines, are attenuated at birth in children who develop atopy subsequently. Most previous studies focused on the relationship between Th1/Th2 cytokines at birth and future development of atopic diseases. However, a few reports examined specifically links between levels of serial cytokines, which play roles in inflammation, maturation of T cells, and production and maintenance of the Th1/Th2 balance, at birth and future development of atopic diseases.

The pathogenesis of AD involves both allergic predisposition and nonallergic environmental factors. Skin barrier disruption has attracted attention as a nonallergic etiologic factor for AD, characterized by disorders of water retention and skin barrier function.8 One study reported that levels of ceramide, a lipid contributing to skin function, were significantly lower in lesional and nonlesional skin of subjects with AD, compared with control subjects.9 However, we know of no reported data concerning when in life dry skin and subsequent skin barrier disruption first become apparent in children who develop AD.

The purpose of the present prospective birth cohort study was to assess the risk of prenatal maternal factors and immunologic profiles at birth and skin functional parameters just after delivery and at 1 month of age in development of AD during the first year of life. To do this, we measured levels of 17 kinds of cytokines, including proinflammatory cytokines, Th1/Th2 distinguishing cytokines, nonspecifically acting cytokines, and chemokines, in serum from cord blood and measured stratum corneum hydration with an impedance meter until 5 days after delivery and again at 1 month.

METHODS

Subjects

This was a prospective cohort study examining multiple prenatal and perinatal factors in relation to child health outcomes. Participants were recruited at Ozawa Obstetric Clinic in Gunma Prefecture, where we explained the study to pregnant women and obtained informed consent at the routine clinic visit corresponding to 35 to 37 weeks of gestation. We enrolled 279 pregnant women (93.9% of eligible subjects) between June 1, 2002, and May 31, 2003. Only children born at term without significant neonatal respiratory difficulties or pathologic jaundice were included in the study, to avoid influences on infant skin physiologic features from humidification in an infant incubator or dehydration during phototherapy. These criteria excluded only 10 children.

Cord blood was collected from participants at birth. Blood was centrifuged at 3000 × g for 15 minutes, and serum was separated and stored at −30°C until cytokine measurement. Mothers were provided a self-administered questionnaire and interviewed briefly. Pregnancy and perinatal data were collected from perinatal records. The main factors taken into account were parental history of atopic disease (asthma, eczema, or hay fever), maternal age, viral infection (upper airway or gastrointestinal infection) or bacterial infection (urinary tract or vaginal infection attributable to Escherichia coli, Chlamydia spp, Gram-positive group B streptococcus, or other bacteria) during the prenatal period, gestational age, infant birth weight, and method of delivery. This study was approved by the committee of ethics at the Department of Pediatrics and Developmental Medicine, Gunma University Graduate School of Medicine.

Follow-up Examinations

In physical examinations at the 1-month checkup visit in our clinic, one of the authors noted the presence or absence of eczema on the face. At that time, physiologic skin measurements also were performed. Children were diagnosed as having infantile eczema when facial eczema was present at 1 month of age. All children were monitored for at least 1 year, for assessment of the development of atopic diseases such as AD, asthma-like illness, and food allergies. Parents were asked directly,
through mail or telephone interviews, whether they had been told by a physician that their child had AD. Information on lifestyle factors and other potential risk factors, including breastfeeding history (exclusive breastfeeding, partial breastfeeding, or milk formula feeding), also was collected with a parental questionnaire.

Skin Physiologic Measurements
Stratum corneum hydration was measured on the forehead, cheek, flexor aspect of the forearm near the cubital fossa, and chest at 5 to 10 hours after delivery and then once daily until 5 days after delivery. All tests were performed in open cribs in a controlled environment, with room temperature ranging from 22°C to 24°C and humidity at 50%. Skin temperature remained stable during the examination for all newborns. Stratum corneum hydration was measured by using a moisture meter (ASA-M1; Asahi Biomed, Tokyo, Japan), based on capacitance and electrical conductance determined at 2 different frequencies (160 Hz and 143 kHz) with 2 concentric surface electrodes. The probe was pressed on the skin surface for 1 to 2 seconds. Each measurement was obtained twice at the same site; data were rejected when children were crying or visibly sweating. Parameters obtained were moisture content on the skin surface and moisture content in the stratum corneum.

Multiplex Cytokine Array Analysis
Multiplex cytokine array analysis was performed by using the Bio-Plex protein array system (Bio-Rad Laboratories, Hercules, CA), using Luminex-based technology. With this assay, we quantitated cytokines simultaneously in serum from cord blood, including IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-12, IL-13, and IL-17, IFN-γ, and TNF-α, granulocyte/macrophage colony-stimulating factor, granulocyte colony-stimulating factor, monocyte chemotactic protein (MCP)-1, and macrophage inflammatory protein (MIP)-1β. Most standard curves ranged between 0.2 pg/mL and 3200 pg/mL. At higher and lower concentrations, standard curves became flat and lost linearity. Lower limits of detection for the assays used were 0.2 pg/mL for all cytokines studied. Samples with undetectable concentrations were assigned a value of 0.1 pg/mL (ie, halfway between 0 and the lower limit of detection for the assay).

Statistical Methods
The relationship between eczema in infancy and various maternal and perinatal risk factors was assessed by using Pearson χ² tests. Differences in clinical characteristics at birth and in skin physiologic parameters between children with and without AD in the first year of life were analyzed by using Student’s t tests. Because the distribution of cytokines is highly skewed, with many values below the lower limit of detection, the Mann-Whitney U test was used to compare cytokine concentrations in cord blood between children with and without AD or infantile eczema. Multivariate logistic regression models were used to determine independent effects of different factors associated with AD in this population; results are expressed as odds ratios (ORs) with 95% confidence intervals (CIs). Subject characteristics and prenatal factors that were statistically significant in Pearson χ² tests, Student’s t tests, or Mann-Whitney U tests were included in the multivariate model.

RESULTS
Family History and Prenatal Factors
Children in this study were monitored prospectively, and questionnaire data from 213 (79%) were collected to assess the incidence of AD during the first year of life. The cumulative incidence of maternally reported, physician-diagnosed AD in the cohort was 27 cases (12.4%) among 213 children. Infant gender, method of delivery, length of gestation, birth weight, and maternal age did not differ between children with and without AD (Table 1). Exclusive breastfeeding was not related to the devel-

**TABLE 1**
Perinatal and Postnatal Characteristics of Infants With AD or Infantile Eczema

<table>
<thead>
<tr>
<th></th>
<th>AD Yes (N = 27)</th>
<th>AD No (N = 186)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/female, n</td>
<td>16/11</td>
<td>86/100</td>
<td>.206</td>
</tr>
<tr>
<td>Cesarean section, n (%)</td>
<td>3 (11.1)</td>
<td>26 (14.0)</td>
<td>.916</td>
</tr>
<tr>
<td>Gestational age, mean ± SD, wk</td>
<td>40.1 ± 12</td>
<td>39.9 ± 11</td>
<td>.343</td>
</tr>
<tr>
<td>Birth weight, mean ± SD, g</td>
<td>3068 ± 310</td>
<td>3121 ± 337</td>
<td>.462</td>
</tr>
<tr>
<td>Maternal age, mean ± SD, y</td>
<td>27.7 ± 30</td>
<td>28.2 ± 43</td>
<td>.534</td>
</tr>
<tr>
<td>Breastfeeding, n (%)</td>
<td>14 (51.9)</td>
<td>102 (54.8)</td>
<td>.771</td>
</tr>
<tr>
<td>Infantile eczema, n (%)</td>
<td>6 (22.2)</td>
<td>20 (10.8)</td>
<td>.089</td>
</tr>
<tr>
<td>Maternal infection, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteriala</td>
<td>5 (18.5)</td>
<td>22 (11.8)</td>
<td>.135</td>
</tr>
<tr>
<td>Viraltb</td>
<td>11 (40.7)</td>
<td>81 (43.5)</td>
<td>.783</td>
</tr>
<tr>
<td></td>
<td>17/9</td>
<td>85/102</td>
<td>.057</td>
</tr>
<tr>
<td></td>
<td>5 (19.2)</td>
<td>121 (11.2)</td>
<td>.396</td>
</tr>
<tr>
<td></td>
<td>39.8 ± 1.1</td>
<td>39.9 ± 1.1</td>
<td>.692</td>
</tr>
<tr>
<td></td>
<td>3119 ± 299</td>
<td>3114 ± 339</td>
<td>.935</td>
</tr>
<tr>
<td></td>
<td>27.7 ± 3.8</td>
<td>28.2 ± 4.2</td>
<td>.551</td>
</tr>
<tr>
<td></td>
<td>15 (50)</td>
<td>103 (55.1)</td>
<td>.626</td>
</tr>
<tr>
<td></td>
<td>17/9</td>
<td>85/102</td>
<td>.057</td>
</tr>
<tr>
<td></td>
<td>5 (19.2)</td>
<td>121 (11.2)</td>
<td>.396</td>
</tr>
<tr>
<td></td>
<td>39.8 ± 1.1</td>
<td>39.9 ± 1.1</td>
<td>.692</td>
</tr>
<tr>
<td></td>
<td>3119 ± 299</td>
<td>3114 ± 339</td>
<td>.935</td>
</tr>
<tr>
<td></td>
<td>27.7 ± 3.8</td>
<td>28.2 ± 4.2</td>
<td>.551</td>
</tr>
<tr>
<td></td>
<td>15 (50)</td>
<td>103 (55.1)</td>
<td>.626</td>
</tr>
</tbody>
</table>

a Urinary tract infection or vaginal infection with Chlamydia spp, Escherichia coli, group B streptococci, or other bacteria.
b Upper airway infection or viral gastroenteritis.
development of AD (Table 1). Rates of maternal bacterial or viral infection during pregnancy did not differ between the 2 groups (Table 1). The diagnosis was >3 times as likely among infants born to mothers with a history of eczema, compared with those born to mothers with no such history. Paternal history of eczema did not influence the likelihood of AD, whereas paternal history of hay fever was associated with less occurrence of AD (Table 2).

The incidence of physician-diagnosed infantile eczema at the age of 1 month was 26 cases (12.1%) among 213 children (Tables 1 and 2). Infantile eczema was more common in male infants than in female infants, although this gender difference was not significant. Infants with infantile eczema were likely to develop AD during the first year (Table 1). Parental history of atopy was not related to the likelihood of infantile eczema (Table 2).

Cytokine Profiles
The proportions of samples with detectable cytokine concentrations were 22% for IL-2, 1% for IL-4, 87% for IL-6, 99% for IL-8, 46% for IL-10, 43% for granulocyte/macrophage colony-stimulating factor, 31% for IFN-γ, 55% for TNF-α, 52% for IL-1β, 36% for IL-5, 91% for IL-7, 16% for IL-12, 30% for IL-13, 48% for IL-17, 70% for granulocyte colony-stimulating factor, 100% for MCP-1, and 100% for MIP-1β. No additional analyses of IL-4 were performed, because of the small number of samples with detectable concentrations. Associations between cord blood cytokines and AD outcomes were analyzed first as a categorical variable (detectable or undetectable) and then in terms of concentrations. There was no difference in the categorical variable for each cytokine between children with and without AD (data not shown). Table 3 compares cytokine concentration profiles in cord blood samples from infants with or without development of AD during their first year. MIP-1β levels were significantly lower in samples from infants who developed AD than in those from infants who did not. A trend toward decreased IL-7 and MCP-1 concentrations in infants with AD fell short of significance (Table 3).

Unlike AD, infants who developed infantile eczema had significantly higher cord blood concentrations of IL-5, IL-17, and MCP-1 than did those who did not (Table 3).

Skin Parameters
The moisture content of the skin surface and that in the stratum corneum were significantly lower on the forehead, cheek, chest, and forearm on postnatal days 2 to 4 than at the age of 1 month (P < .0001) (Table 4). The surface moisture content on the forehead, cheek, and forearm was significantly higher than that on the chest on postnatal days 2 to 4 and at the age of 1 month. The moisture content in the stratum corneum was significantly higher on the cheek than on the forehead, chest, or forearm on postnatal days 2 to 4 (P < .0001) and was significantly higher on the forearm than on the forehead, cheek, or chest at the age of 1 month (P < .0001).

No significant differences in the moisture content of the surface or in the stratum corneum were evident on postnatal days 2 to 4 between infants with and without subsequent AD. In contrast, the moisture content of the surface and stratum corneum of the forehead and cheek at 1 month of age was significantly greater for infants with AD than for those without AD. Infants who developed infantile eczema showed no difference from those who did not in surface or stratum corneum moisture content for any region or time point, except for greater moisture content on the surface of the cheek at the age of 1 month.

Multivariate logistic regression analysis was used to assess independent effects of the various postnatal and prenatal risk factors on development of AD (Table 5). All factors that approached significance with Pearson χ² tests, Student’s t tests, or Mann-Whitney U tests were included in the model, that is, paternal hay fever, maternal eczema, MIP-1β levels, surface moisture content on the forehead and cheek, and moisture content of the stratum corneum of the forehead and cheek, all at the age of 1 month. Paternal hay fever (OR: 0.129; 95% CI: 0.020–0.845), MIP-1β levels (OR: 0.982; 95% CI: 0.967–0.998), and moisture content on the surface of

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Family History for AD and Infantile Eczema</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AD</td>
</tr>
<tr>
<td></td>
<td>Yes (N = 27), n (%)</td>
</tr>
<tr>
<td></td>
<td>No (N = 186), n (%)</td>
</tr>
<tr>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Paternal history</td>
<td></td>
</tr>
<tr>
<td>Eczema</td>
<td>2 (7.4)</td>
</tr>
<tr>
<td></td>
<td>10 (5.4)</td>
</tr>
<tr>
<td></td>
<td>.669</td>
</tr>
<tr>
<td>Asthma</td>
<td>1 (3.7)</td>
</tr>
<tr>
<td></td>
<td>8 (4.3)</td>
</tr>
<tr>
<td></td>
<td>.885</td>
</tr>
<tr>
<td>Hay fever</td>
<td>2 (7.4)</td>
</tr>
<tr>
<td></td>
<td>61 (32.8)</td>
</tr>
<tr>
<td></td>
<td>.007</td>
</tr>
<tr>
<td>Maternal history</td>
<td></td>
</tr>
<tr>
<td>Eczema</td>
<td>4 (14.8)</td>
</tr>
<tr>
<td></td>
<td>8 (4.3)</td>
</tr>
<tr>
<td></td>
<td>.027</td>
</tr>
<tr>
<td>Asthma</td>
<td>1 (3.7)</td>
</tr>
<tr>
<td></td>
<td>9 (4.8)</td>
</tr>
<tr>
<td></td>
<td>.794</td>
</tr>
<tr>
<td>Hay fever</td>
<td>6 (22.2)</td>
</tr>
<tr>
<td></td>
<td>49 (26.3)</td>
</tr>
<tr>
<td></td>
<td>.647</td>
</tr>
</tbody>
</table>
the cheek at an age of 1 month (OR: 3.189; 95% CI: 1.279–7.952), but no other factors, were significant predictors of AD in the best-fitting model.

**DISCUSSION**

Development of atopic diseases is thought to depend on a complex interplay of genetic factors, environmental

---

**TABLE 4** Skin Parameters at 2 Days After Delivery and at 1 Month, According to Development of AD During the First Year and Infantile Eczema During the First Month

<table>
<thead>
<tr>
<th></th>
<th>AD</th>
<th>Infantile Eczema</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes (No)</td>
<td>P</td>
<td>Yes (No)</td>
</tr>
<tr>
<td><strong>Surface moisture, mean ± SD, µS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forehead</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 d</td>
<td>0.47 ± 0.49</td>
<td>0.10 (0.10–0.10)</td>
<td>0.43 ± 0.50</td>
</tr>
<tr>
<td>1 mo</td>
<td>1.88 ± 0.97</td>
<td>1.44 ± 0.53</td>
<td>1.68 ± 0.78</td>
</tr>
<tr>
<td>Cheek</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 d</td>
<td>0.58 ± 0.49</td>
<td>0.10 (0.10–0.10)</td>
<td>0.49 ± 0.28</td>
</tr>
<tr>
<td>1 mo</td>
<td>1.72 ± 0.86</td>
<td>1.24 ± 0.43</td>
<td>1.63 ± 0.78</td>
</tr>
<tr>
<td>Chest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 d</td>
<td>0.30 ± 0.21</td>
<td>0.10 (0.10–0.10)</td>
<td>0.32 ± 0.25</td>
</tr>
<tr>
<td>1 mo</td>
<td>1.07 ± 0.31</td>
<td>1.02 ± 0.30</td>
<td>1.03 ± 0.27</td>
</tr>
<tr>
<td>Forearm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 d</td>
<td>0.57 ± 0.56</td>
<td>0.10 (0.10–0.10)</td>
<td>0.61 ± 0.62</td>
</tr>
<tr>
<td>1 mo</td>
<td>1.54 ± 0.60</td>
<td>1.40 ± 0.50</td>
<td>1.41 ± 0.44</td>
</tr>
<tr>
<td><strong>Stratum corneum moisture, mean ± SD, µS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forehead</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 d</td>
<td>13.5 ± 3.6</td>
<td>14.2 ± 5.8</td>
<td>14.2 ± 5.8</td>
</tr>
<tr>
<td>1 mo</td>
<td>36.2 ± 9.0</td>
<td>32.5 ± 7.8</td>
<td>32.4 ± 9.5</td>
</tr>
<tr>
<td>Cheek</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 d</td>
<td>18.3 ± 4.9</td>
<td>170 ± 6.1</td>
<td>163 ± 4.3</td>
</tr>
<tr>
<td>1 mo</td>
<td>35.6 ± 6.5</td>
<td>31.3 ± 7.1</td>
<td>32.6 ± 5.8</td>
</tr>
<tr>
<td>Chest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 d</td>
<td>13.9 ± 2.7</td>
<td>13.8 ± 3.1</td>
<td>14.2 ± 3.0</td>
</tr>
<tr>
<td>1 mo</td>
<td>29.9 ± 7.2</td>
<td>283 ± 5.7</td>
<td>278 ± 5.9</td>
</tr>
<tr>
<td>Forearm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 d</td>
<td>13.8 ± 5.9</td>
<td>14.7 ± 3.4</td>
<td>16.4 ± 7.5</td>
</tr>
<tr>
<td>1 mo</td>
<td>45.2 ± 11.4</td>
<td>41.3 ± 9.4</td>
<td>43.4 ± 9.9</td>
</tr>
</tbody>
</table>

GM-CSF indicates granulocyte/macrophage colony-stimulating factor; G-CSF, granulocyte colony-stimulating factor.
exposure to allergens, and nonspecific adjuvant factors. Only a few prospective birth cohort studies have addressed AD occurring in the first year of life.6,7 We found several perinatal predictors of increased risk, including certain physiologic skin parameters, low cytokine concentrations in cord blood, and family history of atopic diseases.

We determined the cumulative incidence of maternally reported, physician-diagnosed AD in the cohort of infants studied to be 27 cases (12.4%) among 213 children, in essential agreement with other epidemiologic studies in the United Kingdom12 and the German Multicenter Atopy Study,13 which showed that 14% and 13.4% of subjects, respectively, developed AD in the first year of life. Our study demonstrated increased risk of AD for infants born to mothers with a history of eczema, which is consistent with previous reports that a positive maternal history predicted greater risk for childhood eczema than did a positive paternal history.14 Atopy may be inherited preferentially through the maternal line or mothers may carry relatively more of the predisposing genes. Interestingly, we found a negative relationship between AD and paternal history of hay fever caused by Japanese cedar pollen. Genetic susceptibility to hay fever apparently does not contribute to the development of AD in children; indeed, paternal hay fever might provide a child with some protection against developing AD. Kurzius-Spencer et al12 similarly found evidence of an apparent protective effect of paternal asthma, but mechanisms underlying these protective effects remain unknown.

Environmental influences during pregnancy and early life, particularly those related to hygiene and infections, seem to increase risks of asthma and allergic disease. In fact, certain infectious complications during pregnancy, such as respiratory tract infections, have shown associations with the development of asthma in childhood.15,16 However, we could not detect associations between maternal viral or bacterial infections and the development of AD in children. This disagreement might involve differences in prenatal and postnatal actions of risk factors between various allergic diseases. For example, Calvani et al17 demonstrated that most such factors affected mainly the risk of nonatopic but not atopic asthma. Another possible explanation is that our study was prospective, whereas most others had a cross-sectional design.

In the present study, cord blood concentrations of IL-6, IL-8, IL-10, granulocyte/macrophage colony-stimulating factor, TNF-α, IL-1β, IL-7, IL-17, granulocyte colony-stimulating factor, MCP-1, and MIP-1β were measurable in >40% of samples. However, other cytokines rarely were present in detectable amounts. The reason for the small number of samples with detectable concentrations of these cytokines might be related to suboptimal methods used for their quantification, because assays for some of the cytokines are known to be technically difficult. Alternatively, it is possible that various immune functions, including the capacity to secrete some cytokines, are attenuated at birth.5,18 We could perform no additional analyses of IL-4 because the proportion of samples with detectable concentrations was only 1%.

In the present study, we found negative relationships between cord blood concentrations of MIP-1β and the risk of AD during the first year of life. We know of no previous reports suggesting that MIP-1β within the fetoplacental unit might influence susceptibility to subsequent disease development. The ability of chemokines to regulate Th1 and Th2 responses suggests that these mediators may take part in the pathogenesis of atopic diseases such as allergic asthma, for which Th2 response dominance has been observed. Influences of MIP-1α, MIP-1β, and regulated on activation, normal T cell expressed and secreted in the establishment of Th1 responses have been reported.19 Grob et al20 demonstrated intracellular expression of MIP-1β in CD4+ and CD8+ T cells from patients with allergic asthma to be significantly less than that in cells from subjects without asthma. This observation of diminished MIP-1β production by both CD4+ and CD8+ T cells suggests the relevance of this chemokine to disease development, with relative deficiency being likely to reflect dominance of Th2 responses over Th1 responses at the chemokine level. This view was supported by our present findings showing that low concentrations of cord blood MIP-1β were related to the risk of AD during the first year of life.

Infants with AD showed a trend toward lower concentrations of IL-7 and MCP-1, although the trend did not reach significance. Schonland et al21 demonstrated IL-7 to be a powerful stimulator of neonatal T cells, driving most CD4+ and CD8+ T cells into the cell cycle. Furthermore, the combinations of IL-4 and IL-12 with IL-7 were found to provide superior enhancement of antigen-specific T cell proliferation.22 Although MCP-1 was originally described for its chemotactic activity on monocytes, in vitro studies revealed an even higher activity on T cells.23 Low concentrations of IL-7 and MCP-1 at birth may lead to impairment of T cell activa-
tion; therefore, infants may develop AD later during the first year of life. Another possible explanation for the lower cytokine levels might be secondary phenomena attributable to a different pattern of specific cell subtypes within the blood of these children. Additional studies are needed to confirm whether infants with AD have significantly lower concentrations of IL-7 and MCP-1 at birth.

In 1997, the European Group for Efficacy Measurements on Cosmetics and Other Topical Products gave recommendations regarding electrical measurement methods. According to those recommendations, both single-frequency and multifrequency instruments may be used to assess skin hydration. High-frequency measurements in general reflect the deeper living layers of the skin, whereas low-frequency measurements are dominated by the stratum corneum. In the present study, we used a novel moisture meter (ASA-M1) that measures electrical admittance and susceptance at different excitatory frequencies. We found significant differences between functional skin variables in neonatal life and infancy at the age of 1 month, as well as differences between regions for both neonates and infants. This suggests that stratum corneum function was still adapting to extrauterine life during the period studied. A similar conclusion was drawn when skin surface capacitance and electrical conductance were examined in newborns. Hoeger and Enzmann found a significant increase in stratum corneum hydration, paralleled by a decrease in skin roughness, in serial measurements at 3 days, 4 weeks, and 12 weeks of age.

To our knowledge, no studies have monitored skin function parameters prospectively, to compare directly children with and without subsequent development of AD. In the present study, we found no differences in the moisture content of the surface or stratum corneum, measured a few days after delivery, between infants with and without development of atopy. In contrast, we found significant differences in the moisture content of the surface and stratum corneum on the face at the age of 1 month. Only 6 of 27 infants who developed AD during their first year had facial infantile eczema at 1 month. Although the other 21 infants had no eczema on the face, certain differences in skin function parameters were demonstrated at 1 month between infants with and without subsequent AD. These data suggest that differences in skin physiologic features between infants with and without AD emerge during the first month of life but not in the first few days. We do not know the mechanisms underlying changes in skin function parameters in infants with AD during the first month of life, but abnormalities during this period may be related to impairment of skin adaptation to extrauterine life.

Dry skin, leading to skin barrier disruption, has attracted attention as a nonallergic etiologic factor in AD. Several studies have demonstrated lower water-holding capacity in visually “uninvolved” skin of children with AD, compared with children without AD. Contrary to our expectation, infants with AD had more moisture content in the surface and stratum corneum at the age of 1 month than did infants without AD. The earliest lesions of infantile AD are erythematosus weepy patches on the cheeks, with subsequent extension to the rest of the face and neck. With increasing age, there is a tendency toward drying and thickening of the skin in the involved areas. Therefore, our findings at the age of 1 month may be consistent with the earliest lesions in the clinical course of infantile AD. Furthermore, Yosipovitch et al found a positive relationship between stratum corneum hydration, as evaluated with capacitance measurements, and transepidermal water loss (which reflects skin barrier function in newborns), which indicates that electrical properties of newborn skin may provide an indirect measurement of transepidermal water loss; this was also suggested by Saijo and Tagami and Okah et al. It will be necessary to monitor the skin physiologic features of these infants during the first year, to understand when the findings may eventually change into dry skin, leading to the defective skin barrier that is known in AD.

Infants with infantile eczema at the age of 1 month showed differences in immunologic and skin physiologic parameters, compared with findings for infants with later development of AD during the first year of life. In fact, infants with infantile eczema showed significantly higher concentrations of IL-5, IL-17, and MCP-1. These data suggest that the intrauterine environment is more likely to reflect the development of infantile eczema, rather than that of AD. Recent data indicate that the proinflammatory cytokine IL-17 stimulates the recruitment and activation of neutrophils and macrophages. Furthermore, IL-17 regulates expression of adhesion molecules and chemokines in keratinocytes, which participate actively in skin inflammatory diseases. MCP-1 has been shown to induce the migration of monocytes, which form a significant component of the inflammatory reaction taking place in the skin. Accordingly, higher concentrations of IL-17 and MCP-1 in cord blood of infants who later develop infantile eczema at the age of 1 month may contribute to the enhancement of inflammatory reactions in the skin of these infants. Infants who developed infantile eczema showed no difference from those who did not in stratum corneum moisture content for any region or time point, whereas a significant difference was seen for infants who developed AD during their first year. The reason for the difference in skin physiologic parameters for infants with eczema at the age of 1 month and infants with later development of AD during the first year of life remains unknown. However, measurements of skin physiologic parameters seem to be useful to distinguish infants with later devel-
opment of AD from those with infantile eczema at the age of 1 month.

CONCLUSIONS
Our findings indicated that development of AD in infancy may be related to a decrease in MIP-1β production at birth, greater skin moisture in the surface of the cheek at 1 month of age, and paternal hay fever, in multivariate logistic regression analysis. The majority of risk factors had different effects on infant eczema and AD, which indicates different causes. The number of samples investigated in the present study might be small; larger prospective studies would serve to confirm our results and to explain the possible mechanism of how these factors act.

ACKNOWLEDGMENTS
This work was supported in part by health sciences research grants for research on eye and ear sciences, immunology, allergy, and organ transplantation from the Ministry of Health, Labor, and Welfare of Japan. We thank Tomoko Endo and Chihiro Ijima for their technical assistance.

REFERENCES
Bacterial Imprinting of the Neonatal Immune System: Lessons From Maternal Cells?

Pablo F. Perez, PhD, Joël Doré, PhD, Marion Leclerc, PhD, Florence Levenez, BSc, Jalil Benyacoub, PhD, Patrick Serrant, DESS, Iris Segura-Roggero, MSc, Eduardo J. Schiffrin, MD, Anne Donnet-Hughes, PhD

aNestec, Nestlé Research Centre, Lausanne, Switzerland; bUnit for Ecology and Physiology of the Digestive Tract, National Institute for Agronomic Research, Jouy-en-Josas Cedex, France

Financial Disclosure: Drs Benyacoub, Schiffrin, and Donnet-Hughes, Mr Serrant, and Ms Segura-Roggero were employees of Nestec.

ABSTRACT

OBJECTIVE. We examined the presence of a natural bacterial inoculum in breast milk and its intracellular transport from the maternal intestine to the breast through the circulation.

METHODS. Breast milk and peripheral blood were collected aseptically from healthy donors at various times after delivery, and the presence of viable bacteria was determined through plating. Temporal temperature gradient gel electrophoresis was used to examine the bacterial ribosomal DNA content in milk cells, maternal peripheral blood mononuclear cells, and feces and in corresponding infant feces. Blood from nongravid nonlactating women served as control samples. Bacterial translocation to extraintestinal tissues was also evaluated in virgin, pregnant, and lactating mice.

RESULTS. Breast milk contained a low total concentration of microbes of $<10^3$ colony-forming units per mL. Temporal temperature gradient gel electrophoresis revealed that maternal blood and milk cells contained the genetic material of a greater biodiversity of enteric bacteria. Some bacterial signatures were common to infant feces and to samples of maternal origin. Bacterial translocation from the gut to mesenteric lymph nodes and mammary gland occurred during late pregnancy and lactation in mice.

CONCLUSIONS. Bacterial translocation is a unique physiologic event, which is increased during pregnancy and lactation in rodents. Human breast milk cells contain a limited number of viable bacteria but a range of bacterial DNA signatures, as also found in maternal peripheral blood mononuclear cells. Those peripheral blood mononuclear cells showed greater biodiversity than did peripheral blood mononuclear cells from control women. Taken together, our results suggest that intestinally derived bacterial components are transported to the lactating breast within mononuclear cells. We speculate that this programs the neonatal immune system to recognize specific bacterial molecular patterns and to respond appropriately to pathogens and commensal organisms.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1649
doi:10.1542/peds.2006-1649

Dr Perez’s current affiliation is Centro de Investigación y Desarrollo en Criotecología de Alimentos-Cátedra de Microbiología, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, La Plata, Argentina.

Dr Schiffrin’s current affiliation is Nestlé Nutrition, Nestec, Vevey, Switzerland.

Key Words
bacterial translocation, breast milk, immunity, maternal and child health, lactation

Abbreviations
MLN—mesenteric lymph node
rDNA—ribosomal DNA
TTGE—temporal temperature gradient gel electrophoresis
DC—dendritic cell
PBMC—peripheral blood mononuclear cell
PCR—polymerase chain reaction

Accepted for publication Sep 25, 2006

Address correspondence to Anne Donnet-Hughes, PhD, Nestec, Nestlé Research Centre, Vers-chez-les-Blanc, 1000 Lausanne 26, Switzerland. E-mail: anne.donnet@rdls.nestle.com

PEDEATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275. Copyright © 2007 by the American Academy of Pediatrics

e724 PEREZ et al

at Prince Of Songkla Univ on April 19, 2007 www.pediatrics.org
**Mucosal Dendritic Cells (DCs)**, via pattern recognition receptors such as Toll-like receptors, sample and respond to microbes, which bombard the intestinal mucosa continuously. Normally, this results in tolerance to the normal microbiota and protection against pathogenic attack. Successful simultaneous deployment of such divergent processes requires sophisticated control mechanisms, which are not expected of an inexperienced, neonatal, immune system. However, intestinal colonization and assembly of specific bacterial communities in the absence of adverse immune responses reflect robust regulatory mechanisms, which may already operate in utero. Moreover, differences between breastfed and formula-fed infants in intestinal bacterial colonization and susceptibility to disease suggest that additional regulation is acquired through breast milk.

There is accumulating evidence that bacteria are transmitted to the infant via milk. Most studies of the microbiologic features of milk have addressed the transmission of pathogens or contaminating commensal organisms in samples meant for milk banks. The majority of the latter arise from the mother’s skin or the infant’s mouth. However, certain species are suggested to colonize the neonatal intestine and to provide protection.

The interesting observation that breast milk is not sterile, even when collected aseptically, raises the possibility that breast milk harbors a natural bacterial inoculum, which may influence neonatal colonization. Milk leukocytes are cells that have migrated from the gut- and bronchial-associated lymphoid tissue to lactating mammary glands via the lymphatic vessels and blood circulation. If some microbial species are indeed intrinsic to breast milk, then this cellular circuitry may explain how microbes are conveyed to the breast without any deleterious effect on maternal health. To address this, we examined the presence of bacteria in human milk, blood, and feces during lactation; in a second study, we examined bacterial translocation in nonpregnant, pregnant, and lactating mice.

**METHODS**

**Human Milk, Blood, and Fecal Samples**

Breast milk was collected from healthy lactating mothers who delivered at term. After rejection of ~2 to 3 mL of foremilk, the breast was cleaned with antiseptic soap, rinsed with sterile distilled water, and dried with sterile gauze before aseptic collection with an electrical breast pump. As a control, a swab of the areola was taken before milk collection. Samples of whole milk were plated on de Man, Rogosa, and Sharpe medium containing cysteine, on Eugon tomato, Drigalski, or Shaedler Neo Vanco medium, or on blood agar (bioMérieux, Marcy l’Etoile, France) and were incubated aerobically or anaerobically at 37°C. Leukocytes were collected from the remaining milk through centrifugation and were suspended in sterile phosphate-buffered saline containing 1% gentamicin (10 minutes), to kill extracellular bacteria. Washed cells were then divided into aliquots and were used to make cytopreparations, were frozen in RPMI medium (Life Technologies, Basel, Switzerland) containing 10% dimethylsulfoxide and fetal calf serum for flow cytometric analysis, or were lysed with cold, sterile, distilled water passed through a sterile needle for plating on bacterial culture medium. Bacterial isolates were characterized on the basis of macroscopic and microscopic morphologic features, Gram staining, and culture characteristics.

Approximately 10 mL of venous blood were collected from lactating women at different times after delivery or from 5 age-matched, nongravid, nonlactating women. The blood was centrifuged over Ficoll-Hypaque medium (Sigma-Aldrich, St Louis, MO), washed, and then processed as for milk cells. Maternal and infant fecal samples were collected in sterile tubes, divided into aliquots, and stored frozen at -80°C until required. Written consent was obtained from volunteers, and protocols were approved by our institutional review board and by the Swiss authorities.

**Flow Cytometry**

Myeloid and lymphoid DCs in peripheral blood mononuclear cells (PBMCs) were examined by using the FACSCalibur system and DC-Kit from Becton Dickinson (Basel, Switzerland). In separate tubes, cells were labeled with fluorescein isothiocyanate-anti-CD11c and phycoerythrin-anti-CD14 (Becton Dickinson), according to the manufacturer’s instructions.

**Temporal Temperature Gradient Gel Electrophoresis**

Total DNA was extracted from 200 mg of fecal samples and from milk cells and PBMCs as described previously for feces and biopsies, respectively, except that DNA precipitation of cells was performed overnight and the pellets were centrifuged (1 hour, 4°C). Isolated DNA was then used to amplify the V6–V8 regions of 16S ribosomal DNA (rDNA), with primers U968-GC-F and L1392-R. The polymerase chain reaction (PCR) product size was 468 base pairs. Several dilutions of template DNA were made if the presence of PCR inhibitors was suspected. PCR amplification and temporal temperature gradient gel electrophoresis (TTGE) were performed as reported previously, and Gel Compar II software (Applied Maths, Kortrijk, Belgium) was used to compare TTGE profiles. A PCR amplicon mixture of 7 cloned rDNAs from different bacterial species was used as a migration marker. Some of the TTGE bands that migrated in maternal and infant samples were excised from the gel and sequenced.
Cloning of 16S rDNA

DNA from cells and feces at 4 weeks after delivery were PCR amplified with primers U350-F and L1392-R. The PCR product size was 1080 base pairs. Ligation and cloning were in the pGEM-T vector system (Promega, Madison, WI), except for milk cells, for which 10 PCRs were pooled to make a 16S rDNA library. Forty-eight clones per library were sequenced with the primers M13F and M13R and an equal portion of the SSU rDNA (Escherichia coli positions 350–1392, representing nearly the full-length gene). The sequences from this molecular inventory were longer than those excised from the TTGE gels.

Sequences were checked manually, and the contigs were made by using BioEdit software (Ibis Therapeutics, Carlsbad, CA). The sequences were submitted to GenBank, and the Blast and Megablast programs of the Ribosomal Database Project (East Lansing, MI) were used to identify close phylogenetic relatives. Sequences were tested for chimera structure by using the Ribosomal Database Project analysis service Check Chimera, as well as during manual inspection of alignment. Sequences were compared by using the Blast2sequences program (National Center for Biotechnology Information, Bethesda, MD).

Bacteria Localization in Milk and Blood Cells

After fixation in absolute ethanol, cytopreparations of human milk and blood cells were incubated for 5 minutes with 100 mg/mL acridine orange, washed extensively, mounted in fluorescent mounting medium (Dako Schweiz, Baar, Switzerland), and analyzed with epifluorescence microscopy.

Fetal Liver Tyrosine Kinase-3 Ligand

Fetal liver tyrosine kinase-3 ligand in human serum (one-half dilution) was assayed with an enzyme-linked immunosorbent assay, according to the manufacturer's instructions (R&D Systems, Epalinges, Switzerland). The detection limit of the assay was 10 pg/mL.

Mice

Conventional virgin and pregnant/lactating C57/BL6 mice (Charles River Laboratories, L’Arbresle, France) were killed (n = 10 per group) at 5 to 6 days before parturition or at 1 to 2 days, 3 to 4 days, or 14 to 15 days after parturition. Samples of blood, intestinal contents, mesenteric lymph nodes (MLNs), spleen, liver, and mammary gland were collected aseptically for microbiologic analysis, fixed in Bouin’s fixative before being mounted in paraffin blocks, and/or mounted in OCT medium and frozen in liquid nitrogen. The experimental procedure was approved by our institutional review board and by the Swiss authorities.

Bacteria in Mouse Tissue

Microorganisms were observed in tissue by using Gram stain. For microbiologic analysis, samples of mouse tissue were homogenized, suspended in sterile phosphate-buffered saline, plated onto blood agar (bioMérieux), and incubated aerobically or anaerobically at 37°C.

Statistical Analyses

The proportions of pregnant and lactating animals with viable bacteria in their tissue were compared with that of control animals by using Fisher’s exact test. The median percentages of DC populations in the blood of lactating and control women were compared by using the Mann-Whitney test. TTGE profiles were compared by using Gel Compar II software (Applied Maths). Similarity coefficients (Pearson correlation method) were then calculated for each pair of profiles, yielding a similarity matrix. A dendrogram was constructed from this matrix by using an unweighted pair group method using arithmetic averages algorithm.

RESULTS

Bacterial Signatures Are Transferred From Mother to Infant Through Breast Milk

Skin swabs, made after cleaning the breast with antiseptic soap, did not yield viable bacteria. Aseptically collected breast milk contained a total concentration of <10^3 colony-forming units of bacteria per mL, composed of Lactobacillus, Streptococcus, Enterococcus, Pectostreptococcus, Staphylococcus, Corynebacterium, and/or occasionally Escherichia spp. We next used TTGE to examine bacterial rDNA contents in milk cells and maternal PBMCs during lactation, and we compared the contents with those in maternal and infant fecal samples.

Maternal fecal samples gave classic TTGE profiles that were specific for each individual and of greater biodiversity than those of infant feces (Fig 1A). Although milk cells had a less complex microbiota than maternal feces, TTGE revealed a greater biodiversity than observed previously with plating. Figure 1A shows one mother-infant couple analyzed over weeks 1 to 4 after delivery. Similar TTGE profiles were observed for 6 other mother-infant pairs (data not shown). Interestingly, some bacterial signatures (Fig 1A, arrows) were common in infant feces and in several samples of maternal origin. With excision from the gel and sequencing, the lowest of these milk bands, which was especially intense in infant feces and comigrated in maternal feces and blood, was identified as Bifidobacterium longum on the basis of 369 nucleotides. The presence of B longum was also confirmed in the milk and infant feces of 3 other mother-infant couples (data not shown). One mother also had B longum DNA in her blood cells (Fig 1A), whereas another had the same species in her blood and in her feces. Sequencing of
FIGURE 1
Bacterial signatures in maternal cells and infant feces. Profiling was performed by using TTGE-amplified fragments of rDNA. A, Infant feces, maternal PBMCs, milk cells, and feces (1–4 weeks after delivery). Duplicate PCRs were used for PBMCs and milk cells. A ladder (L) of PCR-amplified, cloned rDNA was used for gel normalization and image analysis. Arrowheads indicate signatures common to infant feces and maternal samples. Excision and sequencing of some bands identified *Bifidobacterium longum* (red arrowheads), *Streptococcus thermophilus/salivarius* (blue arrowheads), and *Staphylococcus epidermidis* (green arrowheads). B, PBMCs from mothers 4 weeks after delivery (lanes 1–5) and control women (lanes A–E). T₀ and Tₚ represent PCR controls. C, Bacterial structures (arrowhead) in milk cells and PBMCs stained with acridine orange.
another band common to milk and infant feces identified DNA from *Streptococcus thermophilus/salivarius*.

Next, PCR products of milk cells were used to prepare rDNA libraries. Besides the species identified previously, sequencing of the clones revealed the presence of *Bacteroides, Clostridium*, and *Eubacterium* among a total of up to 15 genera. Whereas the DNA from staphylococcal and streptococcal species were found in the milk cells of all mothers, DNA from clostridia and lactobacilli were found in the cells of 4 and 3 mothers, respectively. The presence of other genera was specific for each individual. The absence of milk cell genera in the PCR control samples shows that these bacterial DNA were not attributable to laboratory contamination. Lactose-degrading, lactic acid-producing bacteria together with *Staphylococcus* species were the most represented genera in infant feces. Of 23 sequences corresponding to bifidobacteria, 17 were related to *B. longum*, 5 to *Bifidobacterium bifidum*, and 1 to *Bifidobacterium infantis*. Two identical rDNA sequences (99% identity of 1117 base pairs), corresponding to *S. thermophilus* and *Staphylococcus epidermidis*, were identified in the milk cell clones and in the infant’s feces.

PBMCs contained a restricted variety of bacterial rDNA sequences (Fig 1 A and B). Bacterial signals were present in cells of both lactating and nonpregnant nonlactating women, but the complexity of bacterial signatures was greater in the former (Fig 1 B). Furthermore, although profiles for control women were similar, those for lactating women were specific for each individual. Acridine orange staining of milk and blood cytopreparations identified bacterial bodies in association with mononuclear cells (Fig 1 C).

**DC Subsets Are Diminished in the Circulation of Lactating Women**

The distribution of DC phenotypes in the PBMCs of lactating and nonlactating women was examined by using a commercial kit and flow cytometry. The frequencies of differentiated lymphoid DC (lineage -CD14 -HLA-DR +CD11c -CD123 +) and myeloid DC (lineage -CD14 -HLA-DR +CD11c +CD123 +) phenotypes tended to be lower in the circulation of lactating women during the first month after delivery than in that of control subjects (data not shown). This difference reached statistical significance for lymphoid DCs at 1 week after delivery (P = .02) and for myeloid DCs at 3 and 4 weeks after delivery (P = .01 and P = .02, respectively). The numbers of CD14 +CD11c + potential DC precursors were significantly lower throughout the first month after delivery (Fig 2).

**Increased Bacterial Translocation Occurs in Pregnant and Lactating Mice**

Next, bacterial translocation to extraintestinal tissues was examined in conventional nonpregnant, pregnant, and lactating mice. Whereas 10% of control animals had positive MLN cultures, 70% of pregnant animals had bacteria in their MLNs (Fig 3 A). Within 24 hours after delivery, fewer animals had positive MLN cultures but 80% of mice had viable bacteria in their mammary tissue. Although this value decreased to 50% by 3 to 4 days after delivery, it was still significantly different from that of control mice (P < .005). Both aerobic and anaerobic species translocated, and their numbers subsided gradually over time (Fig 3 B).

During lactation, bacteria were observed histologically in the subepithelial dome and interfollicular regions of Peyer’s patches (Fig 3 C, left), in the lamina propria of the small bowel, and associated with cells in the glandular tissue of the mammary gland (Fig 3 C, right). The Peyer’s patches of pregnant and lactating mice were macroscopically larger than those of control animals and had a more prominent subepithelial dome and more dilated draining lymphatic vessels, containing mononuclear cells (Fig 4).

**DISCUSSION**

Aseptically collected breast milk contained a total concentration of microbes of <10^3 colony-forming units per mL, including *Lactobacillus*, *Streptococcus*, *Enterococcus*, *Peptostreptococcus*, *Staphylococcus*, and/or *Corynebacterium*, with occasional *Escherichia* spp. This is less than the concentrations recently reported for breast milk and may reflect elimination of organisms residing in the ducts or on the areola of the breast. Therefore, the findings may give a better indication of bacteria that are intrinsic to milk. It is recognized that, despite every precaution, some of these isolates may still arise from contamination. Several studies have shown a similarity between the microflora of breast milk and that of the...
The organisms most often in common were staphylococcal and streptococcal species. Increases in the number of staphylococci, streptococci, and Lactobacillus acidophilus species after feeding suggest that the infant’s mouth is another potential source of bacteria. Furthermore, a study reported that some strains of Lactobacillus gasseri and Enterococcus faecium in milk were identical to those in swabs of the areola and in oral swabs from the infant. It might be argued that such sources of bacteria are also biologically relevant to neonates. Indeed, Staphylococcus species of the skin are common constituents of the early neonatal microbiota. However, a bacterial presence in all of the milk samples we examined suggests that a discrete microbiota may exist naturally in breast milk. This prompted a subsequent investigation regarding its origin.

We considered that mononuclear phagocytes destined for the mammary gland capture components of the luminal microbiota before their departure from the gut and transfer them to the suckling infant through breast milk. In a first instance, we used TTGE to examine bacterial rDNA content in milk cells and maternal PBMCs and feces during lactation and then examined corresponding infant feces to address transfer of maternal bacteria through milk. Maternal feces yielded classic TTGE profiles that were specific for each mother and of greater biodiversity than those of infant feces. Although milk cells had a less-complex microbiota, TTGE revealed a greater biodiversity than the 2 or 3 genera observed in control animals (C, X 100).
through plating and included genera corresponding to dominant autochthonous ileal and colonic organisms. These results confirmed the expected uptake of bacteria at these tissue sites and suggested that nonculturable bacteria or the DNA from dead bacteria may also be present intracellularly.

Interestingly, PBMCs contained a restricted variety of bacterial rDNA sequences that was more extensive during lactation. No viable bacteria were isolated. The reason for this is unknown, but perhaps the few, bacterially laden cells are diluted in the circulation. Alternatively, bacteria may be dead/quiescent because of intracellular antimicrobial effects.

Migration of bacteria within intestinally derived cells to the breast is supported by the observation that some rDNA bands were common to maternal feces, blood, and milk. Furthermore, because certain of these bands comigrated with those in infant feces, they may represent microbes transferred to the infant through the milk. Indeed, rDNA sequences corresponding to *S thermophilus*, *S epidermidis*, and *B longum* were identified in the milk cells and in the infant’s feces. These 3 species were also detected in other milk samples and infant fecal samples. Moreover, *B longum* was detected in maternal blood and fecal samples.

Because we were aware that PCR amplification might have led inadvertently to false-positive results, we confirmed microscopically whether bacteria were associated with maternal cells. Unlike sepsis, in which translocating bacteria are associated with polymorphonuclear cells,12 bacterial bodies were associated with a limited number (<0.1%) of milk and blood mononuclear cells. However, the possibility that bacterial components are also associated with polymorphonuclear cells cannot be excluded.

The observation that microbial components pass into the circulation of healthy individuals, albeit within an intracellular compartment, is potentially controversial and challenges the dogma that translocation of such material occurs only during sepsis. To verify such a phenomenon, we extended our study to conventional nonpregnant, pregnant, and lactating mice.

Although confined bacterial translocation to the MLNs was seen in control mice, heightened translocation to MLNs in the perinatal period was followed by colonization of the mammary gland in the immediate postpartum period. From this study, we cannot say whether additional “waves” of bacterial translocation occur earlier in pregnancy or later in lactation. Nevertheless, the increased translocation did not seem to be induced solely by parturition. Colonization of the breast coincided with an increased number of positive blood cultures and occasional translocation to the spleen and liver (data not shown). In contrast to pathologic conditions in which translocating microbes are mainly Gram-negative, penetrating species in pregnant and lactating mice included *Streptococcus*, *Lactobacillus*, and *Bifidobacterium*, whose numbers subsided gradually over time.

During lactation, bacteria were observed in the lamina propria of the small bowel and in the subepithelial dome and interfollicular regions of the Peyer’s patches. Therefore, M cell-mediated uptake toward DCs in the Peyer’s patch, direct sampling of luminal bacteria by dendrites of lamina propria DCs, and/or a low-level, physiologic leakiness of the epithelium may occur.14 In healthy animals, a very limited number of bacteria cross the intestinal epithelium, evade uptake and killing by intestinal macrophages, and remain viable after phagocytosis by DCs.15 Bacterially loaded DCs then migrate to the MLNs, where they initiate protective immune responses.15 The more-prominent Peyer’s patches observed in pregnant and lactating animals, with mononuclear cell exit through dilated lymphatic vessels, indirectly suggest that DCs may be implicated in the transport of intestinal microbial components to the breast, through the circuit used for induction of tolerance to soluble antigen.16 Certainly, breast milk has a high proportion of phagocytes, which are also ineffective at killing ingested microbes.17 Work demonstrating that CD14+ milk mononuclear cells, which are normally considered to be macrophages, also express HLA-DR, CD86, CD83, and DC-specific intercellular adhesion molecule-3-grabbing nonintegrin suggests that these cells are partially differentiated DCs.18 Moreover, because tissue macrophages are nonmigrating resident cells, milk DC-like cells derived from the maternal circulation are the most likely vehicles for intestinally derived microbial components. Therefore, we speculated that such cellular populations would be modulated during lactation. Indeed, we found that the frequencies of DC phenotypes and of CD14+CD11c+ intermediate DC-like cells were lower in the circulation during lactation. These findings agree with those of a study showing reduced numbers of circulating DC subsets in late pregnancy19 and may reflect cellular trafficking toward the breast or intestine. We detected fetal liver tyrosine kinase-3 ligand, a stimulator of DC differentiation and mobilization,20 in serum samples of 3 of 9 mothers (range: 13.9–71.7 pg/mL) and in 1 of 5 control samples.

Transfer of bacteria through milk may be a means by which maternal microbes colonize the neonatal gut.5,6 Such a mechanism may provide a colonization advantage to bacteria of the mother’s intestinal microbiota at a time when the low bacterial diversity in the neonatal intestine is permissive to colonization. In the present study, sequence homology between some strains in infant feces and milk suggests that this may indeed occur. However, we observed fewer viable organisms than reported previously5 and, although a greater biodiversity of bacterial DNA was evident in milk cells, not all of those DNA bands comigrated with bands in the infant’s feces. Clearly, there are more efficient routes through which maternal organisms colonize the neonatal gut.
We speculate that this phenomenon represents an education of the neonatal immune system by maternally derived bacterial molecular motifs.

Neonatal immune cells must learn to differentiate between self-antigens, dietary antigens, commensal organisms, and potential pathogens. We showed previously that human milk contains soluble pattern recognition receptors for bacterial motifs and that these may mediate different responses to Gram-negative and Gram-positive organisms and may modulate how neonatal cells perceive and respond to bacterial components. In animal models, uptake of maternal leukocytes into neonatal tissues occurs during gestation and lactation. Perhaps prolonged penetration of inconspicuous bacterial molecular patterns, via maternal DCs during pregnancy and lactation, induces tolerogenic responses that are analogous to those for self-antigens. Interestingly, osteoprotegerin, a DC survival factor that may also be important for maintaining immune tolerance, demonstrates elevated levels in serum during pregnancy and lactation and is present in significant quantities in human breast milk.

Elevated translocation of bacteria or their components in the mother should certainly have some bearing on her immune status and may explain the physiologic activation of innate immunity that occurs during pregnancy. Interestingly, bacterial DNA stimulates innate immunity in pregnant mice, improves maternal survival rates, and prevents pathogen transmission to the fetus.

Our observations suggest a novel form of mother-infant communication, but they also highlight a potentially new mechanism of immune regulation in healthy individuals. As shown previously, the blood of normal healthy subjects contains bacterial components. Some DNA may arise from human or microbial contamination. However, the greater number of bacterial DNA signatures in the PBMCs of healthy lactating women suggests that components of certain bacterial species are inherent to circulating cells. It is tempting to speculate that this represents an evolutionary strategy of immune surveillance and that such bacterial imprinting maintains tolerance to specific bacterial species and alerts distant anatomic sites of changes in local lymphoid tissues.

CONCLUSIONS

Our study shows that human breast milk cells contain a limited number of viable bacteria and bacterial DNA that might have been transported from the mother’s intestine to the mammary gland through an endogenous cellular route. An animal study suggests that this process begins in late pregnancy. The results suggest a novel form of mother-infant communication. However, additional studies are necessary to identify the underlying mechanisms of this heightened bacterial translocation and to elucidate the consequences of this phenomenon for pregnant and lactating women and for instruction of the neonatal immune system.

ACKNOWLEDGMENTS

We thank Fabrizio Arigoni for help in bacterial DNA sequencing; Brigitte Schlosser, Nicole Kusy, Isabelle Rochat, Kim Y. Saudit, Dominique de Maleprade, Angèle Boenzli-Bruand, Paullette Lecoultre, and José-Luis Sánchez for technical assistance; Sylviane Oguey, Anny Blondel, and Ruth Braun for volunteer recruitment and sample collection; and Christine Cherbut, Stephanie Blum, and Irène Cortêsy-Malnoé for scientific discussions and review of the manuscript.

REFERENCES


28. Sacks GP, Redman CW, Sargent IL. Monocytes are primed to produce the Th1 type cytokine IL-12 in normal human pregnancy: an intracellular flow cytometric analysis of peripheral blood mononuclear cells. *Clin Exp Immunol.* 2003;131:490–497


Evaluation of Psychopathological Conditions in Children With Heavy Prenatal Alcohol Exposure

Susanna L. Fryer, MS*, Christie L. McGee, MS*, Georg E. Matt, PhD*, Edward P. Riley, PhDbc, Sarah N. Mattson, PhDbc

*San Diego State University/University of California San Diego Joint Doctoral Program in Clinical Psychology, bDepartment of Psychology, and cCenter for Behavioral Teratology, San Diego State University, San Diego, California

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

ABSTRACT

OBJECTIVE. This study compared the prevalence of psychopathological conditions in children with heavy prenatal alcohol exposure (N = 39) and nonexposed, typically developing peers (N = 30), matched with respect to age, gender, and socioeconomic status.

METHODS. Caregivers were interviewed with either the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children, Present and Lifetime Version, or the Computerized Diagnostic Interview Schedule for Children, Version IV. Statistical resampling methods were used to create 95% confidence intervals for the difference between the proportions of children with psychopathological conditions in the exposed and control groups.

RESULTS. Group differences were seen in the attention-deficit/hyperactivity disorder, depressive disorders, oppositional defiant disorder, conduct disorder, and specific phobia outcome categories. The group difference in the attention-deficit/hyperactivity disorder category was by far the largest effect observed.

CONCLUSIONS. These results suggest that fetal alcohol exposure should be considered a possible factor in the pathogenesis of childhood psychiatric disorders. These data provide clinically relevant information about the mental health problems that children with fetal alcohol exposure are likely to face.
The pattern of birth defects resulting from prenatal alcohol exposure has been studied extensively since the first descriptions of fetal alcohol syndrome (FAS) appeared in the scientific literature.1–3 FAS is associated with a specific clinical presentation, with (1) prenatal and/or postnatal growth deficiency, (2) characteristic facial features (eg, short palpebral fissures, indistinct philtrum, and thin vermillion), and (3) central nervous system dysfunction. Although all 3 indicators must be present to meet the formal diagnostic criteria for FAS, central nervous system dysfunction may occur after prenatal alcohol exposure in the absence of other FAS characteristics.4 In recognition of the wide range of effects after gestational alcohol exposure, the term fetal alcohol spectrum disorders (FASD) is now used as a nondiagnostic umbrella term.5 This new term acknowledges that individuals who fail to meet FAS diagnostic criteria may still show negative effects related to their gestational alcohol exposure. Refining the diagnostic criteria that characterize FASD remains a major research priority.6 The cognitive and behavioral problems associated with fetal alcohol exposure include deficits in learning, language, motor, visuospatial, and executive functioning abilities.4 Deficits in attention and arithmetic skills seem to be especially marked and persistent.7,8 Despite significant cognitive disabilities and lowered IQ, the majority of individuals with prenatal alcohol exposure are not mentally retarded (eg, see ref 9). This fact has important implications, because alcohol-exposed individuals with IQ scores of >70 (ie, those not defined as mentally retarded according to current standards) may not qualify for supportive services, despite evidence that these individuals perform poorly on tests of complex attention, verbal learning, and executive function.10 Adaptive functioning is also affected in this population,9,11,12 and many individuals identified as having FASD are unable to live or work independently.13 Therefore, general cognitive ability alone may not be an effective indicator of special service needs in the FASD population.

Compared with extensive research on cognitive abilities in FASD, mental health outcomes after prenatal alcohol exposure are less well studied. A study of secondary disabilities in FASD revealed that mental health problems had the highest prevalence, compared with other negative outcomes studied (such as disrupted school experience and legal trouble).13 More than 400 individuals with prenatal alcohol exposure were studied, and 94% of the sample experienced mental health problems, according to caretaker reports. Another longitudinal investigation evaluated psychopathological behavior in young children with FAS and observed increased rates of many maladaptive behaviors.14–16 A follow-up report from the same study demonstrated the persistence of psychopathological symptoms through late childhood, including maintained levels of the leading diagnosis of hyperkinetic disorder.17 Although other maternal lifestyle variables, such as smoking, may be important risk factors associated with attention deficits,18 the association between prenatal alcohol exposure and attention deficits is well established in the FASD literature.

There is also evidence that prenatal alcohol exposure is a risk factor for developing depressive features, although alcohol-exposed girls may be more affected than boys.19 Several studies have used Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)-based criteria to evaluate psychopathological conditions in FASD. The Structured Clinical Interview for DSM-IV (SCID) was used to examine psychiatric illness directly in 25 non–mentally retarded, clinically referred, young adults with a history of alcohol exposure.20 Strikingly, 23 (92%) of the 25 participants met SCID criteria for an axis I disorder. Alcohol or drug dependence was the most common diagnosis, followed by major depressive disorder. Similar total rates of psychopathological conditions were found in a mixed sample of nonretarded, 5- to 13-year-old, inpatient and outpatient children with histories of heavy prenatal alcohol exposure.21 Twenty (87%) of 23 children met the criteria for ≥1 of the psychiatric disorders examined, with mood disorders being the most common (including both major depressive and bipolar disorders). More recently, 400 young adults with or without prenatal alcohol exposure were interviewed with the SCID (axis I and axis II versions).22 The study examined whether the high rates of psychiatric illness observed in clinical samples of individuals with FASD would be replicated in a nonclinical, prospectively identified, community sample. The odds of developing substance use disorders (SUDs) and axis II passive-aggressive and antisocial personality traits were twofold greater for alcohol-exposed individuals, after controlling for a multitude of potentially confounding factors, including prenatal nicotine or marijuana exposure, gender, family placement, low socioeconomic status (SES), poor maternal nutrition, and family history of psychiatric problems and alcoholism.

The purpose of the current study was to examine broad-spectrum psychopathological conditions, as defined by the DSM-IV, in a sample of children with heavy prenatal alcohol exposure, compared with nonexposed peers. Increasing awareness among pediatric primary care and mental health providers regarding the psychiatric sequelae associated with fetal alcohol exposure might facilitate the detection of affected children. Such information is especially warranted in this population because early identification and treatment of children affected by prenatal alcohol exposure are associated with improved outcomes.9
METHODS
Study Sample and Procedure
Interviews were conducted with caretakers of 39 children with heavy prenatal alcohol exposure and 30 control children. Alcohol-exposed subjects were drawn from a retrospectively ascertained cohort of >100 children with documented histories of heavy prenatal alcohol exposure (see ref 23 for ascertainment methods). All alcohol-exposed children were evaluated by a dysmorphologist with expertise in the effects of prenatal alcohol exposure (Dr Kenneth Lyons Jones of the University of California, San Diego). The alcohol-exposed group contained children with and without FAS; 15 children (38.5%) had FAS, and 24 children (61.5%) were nondysmorphic alcohol-exposed children. Children in the alcohol-exposed group were all born to mothers who abused alcohol during pregnancy. Although the specific timing and amounts of alcohol exposure are typically unknown under retrospective assessment conditions, histories were determined through a combination of caregiver reports, medical charts, and, if appropriate, social services records. The exposure levels of children studied by our research group were usually in the range associated with DSM-IV-defined criteria for alcohol abuse or dependence. Children with alcohol exposure were referred by Dr Jones, were referred by other medical, social service, or mental health providers, or were self-referred. Children in the comparison groups were self-referred or were recruited through community outreach. The comparison group was matched with the alcohol-exposed group with respect to age, SES, race, and gender (Table 1). Because we were interested in comparing the alcohol-exposed group with a typically developing control group, children were excluded from the comparison group if the following factors were apparent at the time of initial recruitment into the larger research study: exposure to known teratogens, preexisting psychiatric disorders, or parental complaints of notable behavioral problems. With respect to alcohol consumption, mothers of the control children reported no intake or <1 ounce of absolute alcohol intake per day, before pregnancy recognition. Full-scale IQ estimates for children from both groups were based on assessments with the Wechsler Intelligence Scale for Children-III.

| TABLE 1 Sample Demographic Features for Children in the Alcohol-Exposed and Nonexposed Control Groups |
|----------------------------------|----------------------------------|
|                                | Alcohol-Exposed (N = 39) | Control (N = 30) |
| Age, mean ± SD, y              | 12.13 ± 3.16             | 11.20 ± 3.16     |
| Female gender, %               | 46.15                 | 46.67             |
| White race, %                  | 61.54                 | 53.33             |
| SES score, mean ± SD           | 43.91 ± 9.68           | 48.58 ± 10.57    |
| Full-scale IQ, mean ± SD       | 83.03 ± 16.22          | 108.37 ± 12.77*  |

*P < .001.

Psychiatric Interviews
Psychopathological conditions were assessed through standard psychiatric interview measures based on DSM-IV criteria. Two different interviews were used, namely, the Computerized Diagnostic Interview Schedule for Children, Version IV (C-DISC-IV), and the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children, Present and Lifetime Version (K-SADS-PL). The C-DISC-IV is a computer-assisted structured interview developed by the National Institute of Mental Health. The interview version used in this study uses the parent or caregiver as the interview subject. The C-DISC-IV covers axis I DSM-IV disorders, and diagnoses are generated by using an algorithm provided in the software administration package. The K-SADS-PL is a semi-structured psychiatric interview involving separate interviews with the parent or caregiver and child. Diagnoses and impairment ratings based on Diagnostic and Statistical Manual of Mental Disorders, Third Edition, and DSM-IV criteria are generated from answers provided by the interviewee concerning the child’s present and lifetime symptoms.

Interviews performed between 1999 and 2001 (37 interviews) were conducted with the K-SADS-PL by a child psychiatrist who had received specific training on the measure and underwent reliability assessment. Subsequent interviews (32 interviews) were conducted with the C-DISC-IV by psychology graduate students. C-DISC-IV assessments were supervised by a licensed clinical psychologist (Dr Mattson). The measures are equated across groups, such that in the alcohol-exposed sample 17 interviews (43.6%) were conducted with the C-DISC-IV and 22 interviews (56.4%) were conducted with the K-SADS-PL, whereas in the comparison group 15 interviews (50.0%) were conducted with the C-DISC-IV and 15 interviews (50.0%) were conducted with the K-SADS-PL. A previous study comparing clinician-administered K-SADS-P with lay interviewer-administered DISC interviews in an epidemiologic sample demonstrated moderate agreement between the 2 measures.

Statistical Methods
Eleven psychopathological outcome categories were examined, namely, depressive disorders, attention-deficit/hyperactivity disorder (ADHD), oppositional defiant disorder, tic disorders, conduct disorder, generalized anxiety disorder, specific phobia, social phobia, separation...
anxiety disorder, panic disorder, and obsessive-compulsive disorder. Subjects were coded as positive for a particular outcome category on the basis of meeting full criteria on the C-DISC-IV or meeting criteria for a definite past and/or definite present diagnosis on the K-SADS-PL. Between-group analyses were not conducted for the following outcomes because, although assessed, there were no instances of these diagnoses among the study sample: psychotic disorders (e.g., schizophrenia or schizoaffective disorder), bipolar spectrum disorders, post-traumatic stress disorder, eating disorders, agoraphobia, and SUDs. Specific criteria according to psychopathological outcome and psychiatric interview measure are presented in Table 2.

Because of 0-frequency counts in many observational cells, standard, parametric, inferential statistics were not appropriate and logistic regression analysis was not feasible. Instead, data were analyzed by using nonparametric resampling methods (Resampling Stats software; Resampling Stats, Arlington, VA). These computation-intensive methods draw replication samples from observed data to construct confidence intervals based on empirical sampling distributions. Common statistical paradigms that rely on resampling logic include Monte Carlo simulation and bootstrapping techniques. The $\chi^2$ test, with Fisher’s exact test to address comparisons with low cell counts, also could have been used to evaluate between-group comparisons. However, the $\chi^2$ reference distribution is a theoretical probability distribution that is based on parametric assumptions and therefore is less appropriate than the resampling technique, given the small sample size of this study and the nonnormal distribution of the psychopathological outcomes. By using resampling methods, we created 95% confidence intervals for the difference between the proportions of children with psychopathological conditions in the alcohol-exposed group and the matched control group; 95% confidence intervals that contain 0 rule out significant group effects and correspond to a nonsignificant null hypothesis significance test with $\alpha$ set to .05. Alternatively, 95% confidence intervals that do not contain 0 indicate a significant group effect and correspond to a significant null hypothesis significance test at an $\alpha$ level of .05.

RESULTS

In our initial analyses, group differences in factors that may account for variance in mental health outcomes were examined. Specifically, between-group statistical comparisons (alcohol-exposed group versus control group) were conducted for age, SES, and full-scale IQ (independent-sample $t$ test) and for gender, ethnicity, and interview measure ($\chi^2$ analysis). As expected, there was a significant group difference in full-scale IQ ($t_{67} = 7.04; P < .001$); all other matching-variable analyses yielded nonsignificant results. In addition, the groups differed with respect to family placement, with higher rates of nonbiological parent placement in the alcohol-exposed group ($\chi^2 = 31.40; P < .001$). This was expected, because a large proportion of alcohol-exposed children in our sample do not live with biological relatives (26 of 39 nonbiological placements) and the groups were not matched with respect to family placement.

In terms of overall psychiatric diagnoses, 38 (97.44%) of the 39 alcohol-exposed children and 12 (40.00%) of the 30 control children met the criteria for ≥1 axis I disorder. Significant group differences were observed in the following psychopathological outcome categories: ADHD, depressive disorders, oppositional defiant disorder, conduct disorder, and specific phobia. Differences were not seen in tic disorders, generalized anxiety disorder, social phobia, separation anxiety disorder, panic disorder, or obsessive-compulsive disorder (Table 3). Also, there were substantially more observations of comorbid diagnoses in the alcohol-exposed group (27 of 38 cases; 71.05%) than in the control group (6 of 12 cases; 50%). The most common comorbidities occurred among the disruptive disorders (Table 4).

For the alcohol-exposed group, diagnostic differences according to the presence of a FAS diagnosis and family placement were examined qualitatively (because of small sample size) for psychological outcome categories in which significant differences were observed. The pres-
ence of a FAS diagnosis was defined as a binary variable (yes or no). Family placement was also defined as a binary variable, as either biological (living with ≥1 biological parent or biological relative, such as a grandparent or aunt) or nonbiological (foster or adoptive care) placement. In the alcohol-exposed group, 15 children (38.46%) were diagnosed as having FAS, and 26 (66.67%) were living in nonbiological family settings. Qualitatively, slightly higher rates of psychopathological conditions were associated with nonbiological family placement and not having a diagnosis of FAS, although these differences are not likely to be statistically significant, given the small sample size (Table 5).

Lastly, measurement equivalence was examined by looking at the pattern of diagnoses generated with the K-SADS-PL versus the C-DISC-IV within our sample. More specifically, an index score was generated by summing the outcomes (positive diagnosis versus negative diagnosis) for the 5 psychopathological categories in which group differences were observed. Although this type of analysis is less ideal for establishing measurement equivalence than is examination of overlapping data points derived from the 2 types of interviews, the latter type of data was not available for our sample. Therefore, the strategy described above was used to examine measurement comparability, given feasibility constraints. In total, 50 positive and 110 negative diagnoses were generated with the C-DISC-IV, whereas 41 positive and 144 negative diagnoses were generated with the K-SADS-PL. The 95% confidence interval for the proportional difference between the measure-generated outcomes contained 0, providing some evidence that, for the psychopathological outcomes examined, the 2 measures performed with reasonable equivalence in this sample.

TABLE 3 Psychopathological Outcomes and 95% Confidence Intervals for Proportional Group Differences

<table>
<thead>
<tr>
<th>Psychopathological Outcome Category</th>
<th>Alcohol-Exposed (N = 39)</th>
<th>Control (N = 30)</th>
<th>Point Estimate (95% Confidence Interval) for Proportional Difference in Psychiatric Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depressive disorders</td>
<td>7 (17.95)</td>
<td>0 (0.00)</td>
<td>0.18 (0.08 to 0.31)*</td>
</tr>
<tr>
<td>ADHD</td>
<td>37 (94.87)</td>
<td>9 (30.00)</td>
<td>0.65 (0.46 to 0.82)*</td>
</tr>
<tr>
<td>Oppositional defiant disorder</td>
<td>15 (38.46)</td>
<td>5 (16.67)</td>
<td>0.22 (0.02 to 0.42)*</td>
</tr>
<tr>
<td>Tic disorders</td>
<td>3 (7.69)</td>
<td>1 (3.33)</td>
<td>0.04 (−0.06 to 0.15)</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>7 (17.95)</td>
<td>1 (3.33)</td>
<td>0.15 (0.01 to 0.28)*</td>
</tr>
<tr>
<td>Generalized anxiety disorder</td>
<td>3 (7.69)</td>
<td>0 (0.00)</td>
<td>0.08 (0.0 to 0.18)</td>
</tr>
<tr>
<td>Specific phobia</td>
<td>9 (23.08)</td>
<td>1 (3.33)</td>
<td>0.20 (0.05 to 0.35)*</td>
</tr>
<tr>
<td>Social phobia</td>
<td>3 (7.69)</td>
<td>2 (6.67)</td>
<td>0.01 (−0.12 to 0.13)</td>
</tr>
<tr>
<td>Separation anxiety disorder</td>
<td>4 (10.26)</td>
<td>2 (6.67)</td>
<td>0.36 (−0.09 to 0.17)</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>1 (2.56)</td>
<td>0 (0.00)</td>
<td>0.03 (0.0 to 0.08)</td>
</tr>
<tr>
<td>Obsessive-compulsive disorder</td>
<td>1 (2.56)</td>
<td>0 (0.00)</td>
<td>0.03 (0.0 to 0.08)</td>
</tr>
</tbody>
</table>

* Confidence interval does not contain 0, signifying a significant group effect at an α level of .05.

TABLE 4 Numbers of Children With 0 to 7 Diagnoses in the Alcohol-Exposed and Control Groups

<table>
<thead>
<tr>
<th>No. of Positive Diagnoses</th>
<th>Alcohol-Exposed (N = 39)</th>
<th>Control (N = 30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE 5 Qualitative Distributions of Family Placement Status and FAS Diagnosis Among Psychopathological Outcomes in the Alcohol-Exposed Group

<table>
<thead>
<tr>
<th>Psychopathological Outcome</th>
<th>No. of Cases</th>
<th>Family Placement (Nonbiological)</th>
<th>FAS Diagnosis (Positive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADHD</td>
<td>25/37</td>
<td>14/37</td>
<td></td>
</tr>
<tr>
<td>Depressive disorders</td>
<td>5/7</td>
<td>2/7</td>
<td></td>
</tr>
<tr>
<td>Oppositional defiant disorder</td>
<td>11/15</td>
<td>4/15</td>
<td></td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>5/7</td>
<td>0/7</td>
<td></td>
</tr>
<tr>
<td>Specific phobia</td>
<td>6/8</td>
<td>2/9</td>
<td></td>
</tr>
</tbody>
</table>

Denominators represent the number of cases with positive diagnoses in the respective outcome category.

DISCUSSION

Children with heavy prenatal alcohol exposure experienced higher rates of many common DSM-IV axis I psychiatric disorders, compared with matched comparison children. High rates of axis I disorders were documented previously in samples of children and young adults with prenatal alcohol exposure. These data support those findings and extend them by comparing alcohol-exposed children directly with their typically developing peers.

The largest group effect observed in our sample was in the ADHD category, which is not surprising, given the well-documented association between FASD and attention-deficit/hyperactivity disorder (ADHD). The presence of ADHD, along with other severe psychopathological conditions, may place children at higher risk of experiencing adverse social, family, and educational outcomes throughout their lifetimes.
tion deficits. This finding is consistent with previous research focused on fetal alcohol exposure and psychopathological outcomes in children. Evidence from both anecdotal description and empirical studies corroborates the widespread presence of attention problems in the FASD population. In particular, studies have shown that the attention deficits associated with prenatal alcohol exposure occur independently of general intelligence deficits and are persistent and stable over time. A study from our laboratory suggested that children with heavy prenatal alcohol exposure could be distinguished from nonexposed peers on the basis of scores for frequently used measures of attention; alcohol-exposed and comparison subjects were classified with 91.7% accuracy. Therefore, attention deficits, as indicated for the current sample, are well described in the scientific literature on fetal alcohol effects. Although these deficits obviously are not pathognomonic for FASD, data suggest that their presence can be of diagnostic utility. Notably, our control group exhibited an unexpectedly high rate of ADHD (30%). One explanation for this phenomenon is ascertainment bias. Although children were excluded from the comparison group on the basis of premorbid psychiatric conditions or noted behavioral problems, it is possible that the recruited control group still possessed a higher rate of undiagnosed attention problems than found for the population at large. Because participation offered neuropsychological services to the subject at no cost, there might have been some enticement for participation. Another possibility is that our study protocol (ie, diagnosis based on psychiatric interview results, rather than an experienced clinician’s opinion) overdiagnosed the disorder. However, the ADHD between-group effect was quite large (95% vs 30%), which emphasizes that alcohol-exposed children are far more likely to present with ADHD than are their nonexposed peers.

In addition to ADHD, significant group differences were observed within the other disruptive disorder categories, that is, oppositional defiant disorder and conduct disorder. This finding too is reflected in the literature on FASD. Increased delinquent behavior has been reported for alcohol-exposed youths, although factors such as level of exposure and home placement (biological, foster, or adoptive) likely moderate the relationship between prenatal alcohol exposure and delinquency. A qualitative examination of these variables in our alcohol-exposed sample did not suggest different patterns of psychopathological conditions based on family placement or FASD diagnosis; however, this might be attributable to the small sample size. In any case, individuals with prenatal alcohol exposure are thought to be significantly overrepresented in the criminal justice system, although corrections staff members are largely unaware of this phenomenon. One of the few systematic FASD screens of a delinquent group was undertaken in a forensic psychiatric facility in British Columbia. The study revealed that 23% of juvenile detainees were exposed to significant amounts of alcohol prenatally; the majority of these exposures were undetected before the screen. Unlike studies of young adults, which have associated FASD with increased rates of SUDs, we did not observe any instances of SUDs in our sample, although this is likely a reflection of the relative youth of our subjects.

A greater proportion of depressive disorders was also observed in our alcohol-exposed sample, compared with control subjects, which emphasizes that the impairments associated with alcohol teratogenesis are not limited to disruptive behavior. This is an important point, because the internalizing nature of depressive disorders may be more difficult to recognize than externalizing behavior problems, which are observed commonly in the disruptive disorders. The potential link between fetal alcohol exposure and depression has been investigated less frequently than the relationship between FASD and disruptive psychopathological conditions. However, our finding is consistent with research associating prenatal alcohol exposure with the development of negative infant affect and, subsequently, depressive features in children. Also, research using a psychosocial inventory indicated that, as reported by their caregivers, alcohol-exposed children displayed higher rates of depressive features, compared with typically developing children, although factors such as SES may serve as important moderators of the relationship between prenatal alcohol exposure and parental reports of internalizing behaviors.

In our sample, there were no instances of disorders on the bipolar spectrum. This finding diverges from results of a study by O’Connor et al, which documented high rates of bipolar disorder in children with FASD (8 of 23 children: 35% of the sample). One possible reason for this discrepancy might be the inclusion of inpatients in the study by O’Connor et al whereas the present study sampled noninpatients exclusively. It is possible that more severe or different types of psychopathological conditions are associated with inpatient individuals with FASD. Such a rationale may also explain the much higher rates of ADHD in the present study, compared with the study by O’Connor et al (95% vs 13%). Interestingly, both studies failed to observe any instances of psychotic disorders; however, as with SUDs, it is possible that this finding is related to the typical age of onset, because there is a suggestion that increased rates of psychosis occur in alcohol-exposed young adults.

These data indicate that children exposed to alcohol prenatally may present with significant psychopathological conditions. Moreover, certain psychiatric disorders may be more prevalent than others in FASD, because disruptive disorders were particularly common in our sample, whereas many anxiety disorders were not. Ulti-
mately, these data underscore the need for special services, including psychological evaluations, for individuals with histories of prenatal alcohol exposure.

This study should be of interest to clinicians, because its main purpose was to ascertain the frequency and nature of mental illness in a sample of children with documented histories of heavy fetal alcohol exposure. Medically defined diagnoses are a useful step in translating behavioral symptoms of children with prenatal alcohol exposure into clinical relevance and are warranted in the interest of guiding strategies for remediation in this special-needs population. The effects observed in this study support existing literature that suggests that prenatal alcohol exposure is associated with clinically significant psychopathological conditions. Surprisingly, although 97% of alcohol-exposed children had ≥1 axis I disorder, only 40% of our alcohol-exposed sample had been evaluated for or received a psychiatric diagnosis previously. It is hoped that these data will serve to increase awareness among pediatric and mental health caregivers, because many individuals who suffer deleterious consequences from prenatal alcohol exposure remain undetected or misdiagnosed.

This study is limited by several factors. The first is the small sample size (N = 69), which limits the power to detect relationships with small effect sizes. Therefore, the negative findings presented in this study should be interpreted cautiously until they are replicated with a larger sample. Although the modest sample size is a limitation, heavy prenatal alcohol exposure has a low base rate in the population at large, and the small sample size must be weighed against the relative incidence of this condition. In addition to sample size concerns, the psychiatric diagnoses in our study were generated with protocols from psychiatric-based interview tools. Although such measures can be a useful component of an integrated psychological assessment approach, they are not intended to be diagnostically conclusive. In the context of this study, the psychopathological assessment methods used are best viewed as an approximation of an evaluation performed more properly by a trained clinician with expertise regarding the diagnosis in question, who synthesizes data from a variety of sources. However, the psychiatric interview-based assessment approach used was adequate for the study aim, and similar interview-based approaches have been used in studies of other populations.41,42 Also, this study combined interview results from a structured measure (C-DISC-IV) and a semi-structured measure (K-SADS-PL). Although this approach is not ideal, because the measurement equivalence of these 2 methods has not been researched thoroughly, threats to construct validity posed by this limitation were minimized by ensuring that the alcohol-exposed and comparison groups were equally represented by both measures and by examining statistically the patterns of diagnoses (positive versus negative) through interview measurements within our sample.

Another limitation exists because of the retrospective ascertainment of the study sample. Because most individuals from the alcohol-exposed group were clinically referred, it is possible that this study overestimates the association of prenatal alcohol exposure and psychopathological conditions. Therefore, these study results may not be generalizable to the alcohol-exposed population at large. However, this study should inform intervention efforts, and this sample reflects adequately the portion of the alcohol-exposed population that would most likely seek out and benefit from mental health treatment. In addition, previous studies indicated that our alcohol-exposed sample was similar, with respect to several important characteristics (such as IQ and attention abilities), to other samples of individuals with FASD.4 Lastly, the quasi-experimental design of this study involves the comparison of 2 groups that are not equivalent with respect to 2 important factors, namely, family placement and general intelligence, both of which are potential sources of variance in mental health outcomes.43,44 Although these factors possibly confound interpretation of these results, we do not think that the matching scheme used in this study detracts from the overall study aim, to characterize the likely psychopathological presentations of alcohol-exposed children. Future studies of psychopathological conditions and prenatal alcohol exposure should seek to replicate these results by using a comparison sample of mental age- and home placement-matched individuals, to better pinpoint the cause of the FASD-associated psychiatric problems described in this report. Despite these limitations, this study provides important information on the extent of mental health issues in the FASD population and will likely be of relevance to a wide-ranging readership, including health care providers, social service advocates, and individuals suffering negative effects of prenatal alcohol exposure.

ACKNOWLEDGMENTS
This research was funded by National Institute on Alcohol Abuse and Alcoholism grants R01 AA010417, R01 AA012596, R01 AA010820, and F31 AA016051.

We acknowledge the contributions of Drs Brigitte Robertson and Jeffrey Max and the assistance and support of the Center for Behavioral Teratology.

REFERENCES


Di-(2-ethylhexyl)phthalate and Deep Venous Thrombosis in Children: A Clinical and Experimental Analysis

Dirk Danschutter, CCRN, CP, MSc, Filip Braet, PhD, Elke Van Gyseghem, PhD, Said Hachimi-Idrissi, MD, PhD, FCCM, Benny Van Bruwaene, MD, Pat Moloney-Harmon, RN, MS, CCNS, CCRN, FAAN, Luc Huyghens, MD, PhD, FCCM

*Department of Intensive Care, †Department of Medical Registration and Statistics, Free University of Brussels, Brussels, Belgium; ‡Laboratory of Pharmaceutical Technology and Biopharmacy, Catholic University of Leuven, Leuven, Belgium; †Australian Key Centre for Microscopy and Microanalysis, University of Sydney, Sydney, Australia; ‡Advanced Nursing Practice, Children’s Services, Sinai Hospital, Baltimore, Maryland

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

ABSTRACT

BACKGROUND. Five children with catheter-related deep venous thrombosis were encountered in our PICU. Three types of polyvinyl chloride tubing for the administration of intravenous solutions were in use (Terumo, Codan, and Per fusend). All were di-(2-ethylhexyl)phthalate plasticized. We suspected problems with the Codan tubing.

METHODS. Different types of tubing at different time intervals in vitro were investigated. Tubing segments were assessed on structural alterations by surface electron microscopy. High-performance liquid chromatography-diode array detection and liquid chromatography-mass spectrometry-diode array detection were performed to identify and to quantify di-(2-ethylhexyl)phthalate. The hospital’s minimal clinical data set (coded with the International Classification of Diseases, Ninth Revision, Clinical Modification) was investigated on catheter-related deep venous thrombosis between 2000 and 2004.

RESULTS. Surface electron microscopy demonstrated that the Codan tubing’s inner surface was severely altered, showing large particles (34.5 ± 6.1 μm). High-performance liquid chromatography documented that all Codan samples showed a peak at the di-(2-ethylhexyl)phthalate retention time. The analysis of the minimal clinical data set for total catheter-related deep venous thrombosis showed an unusual high incidence in 2001 (52) compared with the expected 36 per year.

CONCLUSIONS. Such occurrence of catheter-related deep venous thrombosis led to the assumption that disintegration of intravenous tubing resulted in intravenous administration of debris. Our data suggested that the particles derived from the tubing are of such size that they might induce catheter-related deep venous thrombosis. The absence of catheter-related deep venous thrombosis caused by the introduction of submicron inline filters outlines the important pathophysiological role of di-(2-ethylhexyl)phthalate-plasticized particles in the onset of catheter-related deep venous thrombosis. Our data indicate that a considerable number of patients might have been exposed to di-(2-ethylhexyl)phthalate, and a major concern is whether this jeopardized the health of the patients at that time.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2221
doi:10.1542/peds.2006-2221

Key Words
di-(2-ethylhexyl)phthalate

Abbreviations
DVT—deep venous thrombosis
CR-DVT—catheter-related deep venous thrombosis
PVC—polyvinyl chloride
DEHP—di-(2-ethylhexyl)phthalate
PF—peristaltic finger
TOTM—trioctyl trimellitate
HPLC-DAD—high-performance liquid chromatography-diode array detection
LC-MS-DAD—liquid chromatography-mass spectrometry-diode array detection
m/z—mass to charge ratio
SEM—surface electron microscopy
HPLC—high-performance liquid chromatography
MS—mass spectrometry
mAbs—milliabsorbance

Accepted for publication Sep 26, 2006
Address correspondence to Dirk Danschutter, RN, CCRN, CP, ANP, MSc, Department of Intensive Care, Free University of Brussels, Laarbeeklaan 101, BE 1090 Brussels, Belgium.
E-mail: dirk.danschutter@vub.ac.be

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
DEEP VENOUS THROMBOSIS (DVT) is a seldom encountered but potentially dangerous medical condition in children. Factors that may predispose to DVT include a hypercoagulable state, phlebitis, systemic or catheter-related infections, irritation, and drug incompatibility.

Over a 33-day period in 2001, we encountered in our PICU 5 children with catheter-related DVT (CR-DVT) of unknown origin. Having excluded primary medical conditions as a possible cause, we extensively investigated the possible etiologic role of fluid and drug incompatibilities and the intravenous administration tubing for intravenous lines. The type of femoral catheters used by the PICU historically consisted of polyethylene and polyurethane. Before the occurrence of the 5 CR-DVT cases, the hospital and subsequently the PICU had gradually begun introducing new intravenous administration tubing sets, as well as new peristaltic (volumetric) intravenous pumps. In all, 2 types of peristaltic pumps (Terumo [Leuven, Belgium], type STC-503 and Argus [Heimberg, Switzerland], type 414) and 3 types of intravenous-sets were in use (Terumo, Codan [Lensahn, Germany], and Perfusend [Sendal, Almaraz, Spain]) during the occurrence period. All intravenous administration tubing sets were commercially manufactured from polyvinyl chloride (PVC) plasticized with di-(2-ethylhexyl)phthalate (DEHP). The intravenous tubing system was pressure driven by peristaltic finger (PF) pumps. With 1 of the tubing systems recently introduced (Codan tubing with Terumo PF pumps), we noticed that the dialed-in volume on the infusion pump ran in over a shorter period of time than programmed. These PF pumps were thoroughly checked and found to be functioning correctly. Hence, we hypothesized that there might be a tubing-related problem.

This made the PICU unit eager to investigate whether tubing wear with potential release of plastic particles and DEHP may have played a role in the etiology of CR-DVT in these patients. We decided to investigate 3 types of tubing used in our PICU: Terumo, Codan V86-P-Y ref. 16/03/14054, and Perfusend. From a prospective interest, we included in our study the intravenous tubing of Cardinalhealth (Hampshire, United Kingdom) for Cardinalhealth type 571 PF pumps.

PATIENTS AND METHODS

Patients

All 5 patients were admitted to a tertiary pediatric care center and required central line infusions. Using a Seldinger technique, the catheters were introduced in the right or left femoral vein by using either a Vygon (Ecoven, France) single-lumen catheter or a Cook (Bloomington, IN) double-lumen catheter. All patients had additional peripheral inserted catheters and an arterial line. The central access served for the administration of inotropes and total parenteral nutrition. As a routine, the compatibility of solutions infused was checked for each patient against the recommendations of the Handbook on Injectable Drugs.

Tubing

We investigated the 3 types of tubing used in the PICU (and through the entire hospital): Terumo, Codan, and Perfusend. From a prospective interest the Cardinal-health tubing used in the NICU was investigated by surface electron microscopy (SEM) alone. The investigators were blind to the type of tubing being investigated.

To assess the accuracy of the different types of PF pumps, a 1-liter bag was connected to Codan, Terumo, and Perfusend tubing and driven by Terumo for Codan and Terumo and Argus PF pumps for Perfusend. The reference solution infused in the tubing during the experiment consisted of 0.9% sodium chloride in water for injection (Baxter, Lessines, Belgium). The volume of solution, administered over 24 hours for all tubing, was calculated to match an infusion for a child weighing approximately 10 kg and amounted to 42 mL per hour. The experiment was allowed to run for 96 hours.

During another experiment, paired samples of 10 mL of fluid (sample “Cod1” and “Cod2”) were collected at 0, 24, 48, 72, and 96 hours from 2 different Codan tubings. The 2 Codan samples collected at 96 hours from 2 different Codan tubings first passed an inline screen filter with a pore size of 0.2 μm. A sample of 10 mL of fluid was collected at 0, 24, 48, and 72 hours from the Perfusend tubing. The distal end of the different tubing was allowed to drain freely into a collector without contact with the latter.

The collected samples were used for identification and quantification of DEHP or triocyl trimellitate (TOTM) by using high-performance liquid chromatography-diode array detection (HPLC-DAD) and liquid chromatography-mass spectrometry-diode array detection (LC-MS-DAD).

Segments from all types of tubing (Codan, Terumo, Perfusend, and Cardinalhealth) in contact with the PF pumps were removed at 0 and 72 hours for SEM investigation. After the initial SEM results of these segments, additional pump segments of the Codan tubing were removed at 6, 12, and 18 hours for SEM investigation.

Surface Electron Microscopy

All types of intravenous tubing pump segments (Codan, Terumo, Perfusend, and Cardinalhealth) were carefully sliced with a razor blade to reveal the inner surface of the tubing. Samples were transferred to a desiccator at room temperature for 24 hours to avoid water contamination. Subsequently, the tubes were mounted on 1-in aluminum pin stubs (G399, Agar Scientific, Essex, United Kingdom) with the aid of double-sided tape, and then sputter coated with 10-nm gold. The specimens were examined with a Philips (Eindhoven, Netherlands) SEM.
505 at an accelerating voltage of 30 kV and a spot size of 20 nm to obtain images with low noise content, as previously described. The magnification of the SEM was regularly calibrated with a cross-grating replica (Polaron, Watford, United Kingdom; 30 000 lines/in) with the specimen in eucentric position. The obtained images were transferred to the UTHSCSA Image Tool 2.0 software (Department of Dental Diagnostic Science, University of Texas Health Science Center, San Antonio, TX) and was used to determine the particle size and for figure assembly.

HPLC-DAD and LC-MS-DAD Investigation of the Presence of DEHP in Collection Fluids

Reagents
Navigated by SEM investigation, solution samples from Codan and Perfusend were analyzed for the presence of either DEHP or TOTM with HPLC-DAD. To avoid environmental contamination, the samples were collected in borosilicate glass tubes with phthalate-free stopcocks. Two more Codan samples from the 2 different Codan tubings were collected after 96 hours runtime while having crossed an inline screen filter with a pore size of 0.2 μm.

A saturated solution of bis-(2-ethylhexyl)phthalate for synthesis (Merck-Schuchardt, Schuchardt, Hohenbrun bei München, Germany) in Milli-Q water was used to obtain the ultraviolet spectrum of DEHP. Milli-Q water was prepared with the Millipore purification system (Millipore, Molsheim, France). In addition, a solution of bis-(2-ethylhexyl)phthalate for synthesis (Merck-Schuchardt) was prepared in acetonitrile, Hypersolv for high-performance liquid chromatography (HPLC; BDH, Poole, United Kingdom), because the solubility of DEHP is higher in the organic solvent. For the same reason, the tubing was immersed for 2 hours in acetonitrile.

For the HPLC-DAD experiments, the solutions and mobile phases were prepared by using acetonitrile, Hypersolv for high-performance liquid chromatography (HPLC; BDH) and Milli-Q water.

To be sure of the nature of the component behind the retention time, additional structural characterization was necessary. Therefore, the experiments were repeated by using LC-MS-DAD, so that through mass spectrometry (MS) the mass to charge ratio (m/z) of the compound could also be determined. The mobile phase of LC-MS-DAD was prepared by using acetonitrile, for HPLC far UV (Acros Organics, Geel, Belgium).

Instruments
The ultraviolet spectrum was recorded ranging from 200 to 400 nm by using a Perkin Elmer-Lambda 20 UV/VIS spectrophotometer (Perkin Elmer, Norwalk, CT). A Bransonic 5210E-MT (Branson Ultrasonic Cooperation, Danbury, CT) ultrasonic bath was used for degassing.

The HPLC-DAD instrument consisted of a model 5000 liquid chromatograph pump (Varian, Palo Alto, CA), a 100-μL loop, a CTO-10A column oven, and an SPD-M10A diode array detector (both Shimadzu, Kyoto, Japan). The chromatographic methods and data were created and treated with the Class-M10A LC workstation software (Shimadzu). The diode array detector scanned in the range of 200 to 400 nm.

The stationary phase used was Discovery RP-AmideC16 (100 × 4.6 mm inner diameter; mesh size: 5 μm; Supelco, Bellefonte, PA). The column can withstand 100% aqueous conditions without showing phase collapse because it is polar-embedded. The flow rate was 1.0 mL/min. The temperature of the oven was kept at 40°C.

Additional experiments were performed by using an LC-MS-DAD instrument consisting of a Waters 2695 separations module (= alliance) HPLC compartment (Waters, Milford, MA), a Mistral column oven (Spark Holland Instrumenten, Emmen, the Netherlands), a column switcher (VICI AG, Schenkon, Switzerland), a Waters 996 Photodiode Array Detector (Waters, Milford, MA), and a quadrupole time-of-flight mass spectrometer (Waters, Manchester, England, United Kingdom). The mass spectrometer data were obtained by using a scan range of 80 to 1000 Dalton (m/z), a scan time of 1 second, a resolution of ± 9000, the multiple channel plate at 2.2 kV, applying electron spray ionization in the positive ion mode as mass spectrometer parameters, and the source temperature of the ion source at 100°C. The chromatographic methods were created and the data treated by using both Millennium 4.0 software (Waters) for the spectral data and MassLynx 3.5 software (Micromass, Cary, NC) for the MS data.

The solution samples were reprocessed by using a linear gradient at a flow rate of 1.0 mL/min. All experiments were conducted at 40°C. The injection volume was 5 μL. The experiments were run on a Hypersil C18-BDS (100 × 4.0 mm inner diameter; mesh size: 3 μm) and a Zorbax SB-C8 (150 × 4.6 mm inner diameter; mesh size: 5 μm) stationary phase, both from Agilent (Palo Alto, CA).

Minimal Clinical Data Set
In Belgium, as in most of the European countries, all diagnoses that patients present during their hospital stay are coded with the International Classification of Diseases. In Belgium, the International Classification of Diseases, Ninth Revision, Clinical Modification, is used.

In the database of the University Hospital of Free University of Brussels, all admissions between 2000 and 2004 with DVT of the femoral vein (code 451.11) as secondary diagnosis were selected and counted in the nominator. The number of patients of the diagnosis-related groups in which at least 1 DVT was noticed was used as denominator. The incidence of CR-DVT was
calculated for the PICU, for the ICUs, and for hospitalization wards for adult patients. Because the Department of Neonatology did not use the Codan tubes, this department was excluded.

RESULTS

Patients
In the PICU, 5 of 117 patients presented with a CR-DVT in 2001, semester 1. Patient 1 (1.25 years old) was admitted after cardiac arrest after accidental strangulation. Patient 2 (4 years old) was hospitalized for encephalitis and deep comatose condition caused by Epstein-Barr viral infection. Patient 3 (1.75 years old) with multiple congenital abnormalities and cerebral cysts was admitted for respiratory syncytial virus pneumonia and respiratory failure attributable to status epilepticus. Patient 4 (15 years old) was admitted for respiratory failure attributable to Duchenne myodystrophy and was transferred to the PICU for postoperative and septic shock after major surgery on the respiratory system. Patient 5 (0.75 years old) was admitted for respiratory failure complicating respiratory syncytial virus pneumonia. Demographic and clinical data of the 4 patients are presented in Table 1. The coagulation profile of the 5 patients is shown in Table 2. The type of solutions infused before onset of the CR-DVT is shown in Table 3. The site of venous thrombosis was established by using echodoppler ultrasound evaluation and found to be at the tip of the catheter and/or short distance upstream. In patient 4 (Table 1), a massive CR-DVT was asymptomatic and subsequently diagnosed by a computed tomography scan of the pelvic region for another condition (acute abdominal pain). Immediately after the diagnosis of thrombosis, all 5 patients were treated with low molecular weight heparin. Two were additionally treated with coumarin (Table 3).

Because of the fifth incidence of CR-DVT at the PICU (patient 5), inline screen filters with a pore size of 0.2 µm were introduced with not a single occurrence of CR-DVT observed in the PICU until June 2006.

Bedside nurses throughout the hospital occasionally observed and reported that Terumo PF pumps did not administer intravenous fluids as accurately as in the past, whereas service engineers were unable to detect any malfunctioning of the equipment.

Surface Electron Microscopy
SEM examination of the Perfusend tubes at 0 hours showed multiple parallel-organized, worm-like structures, which seemed to be packed into multiple layers on top of each other. In contrast, the same Perfusend tubes exposed at 72 hours to the PF pump mechanism showed a more flattened outlook, ie, the worm-like impressions showed a strong parallel organization and multiple layers seemed to be absent. In addition, these tubes showed large surface indentations that were positioned with a periodicity of ~0.14 ± 0.1 mm (n = 20), indicating that these topology changes are derived from the segments of the PF pump-driving mechanism. These indentations could already be observed by macroscopic visual inspection.

Codan tubes at 0 hours revealed smaller worm-like impressions with a random organization (Fig 1) compared with Perfusend tubes at 0 hours (Fig 2). Cardinal-health and Terumo control tubes were revealed to have a rather smooth surface with no worm-like structure.

Codan tubes that were placed in the PF pump mechanism showed a dramatic change in surface topology, ie, large and deep grooves, as early as 6 hours after pumping. These large artificial grooves or cracks gave the impression that the polymer’s surface had lost its integrity and was severely damaged (Fig 3). In addition to these artificial grooves, the surface lacked the small worm-like structures seen at 0 hours and instead large ribbon-like surface corrugations appeared. Moreover, large particles (34.5 ± 6.1 µm [n = 20]) became apparent (Fig 4); these seemed to be fragments of the Codan tube.

Particles with a smaller size (4.5 ± 1.2 µm [n = 20]) could also be noticed on the Terumo tubes (Fig 5) exposed to the PF pump mechanism, although they were present in a lower quantity. Like the Terumo tubes, the Perfusend (Fig 6) and Cardinalhealth PF pump segments kept their wall integrity, even after 96 hours of pumping time.

HPLC-DAD and LC-MS-DAD
The ultraviolet spectrum of DEHP was recorded for a saturated solution of the substance in water to investigate whether the compound was exhibiting a large extinction coefficient. It seemed to have a very high absorbance, which implies that the presence of DEHP in

<table>
<thead>
<tr>
<th>TABLE 1 Patient Demographic and Clinical Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient No.</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

LFV indicates left femoral vein; CVC, central venous catheter; CDL, Cook double lumen; VSL, Vygon single lumen; EBV, Epstein-Barr virus; DMD, Duchenne muscular dystrophy.
the samples would be easily detected (figure not shown). The spectrum shows a wavelength of maximal absorbance at 274 nm.

Injection of the solution of DEHP in acetonitrile applying HPLC-DAD using 95%/5% (vol/vol) acetonitrile/water isocratically as the mobile phase gave rise to a peak at 1.712 minutes/1.691 minutes (220 nm/254 and 274 nm) retention time. Knowing now the retention time of the DEHP, as well as its ultraviolet spectrum in both water and acetonitrile, 2 different Perfusend samples collected at 72 hours from 2 different PF pumps were injected using the same chromatographic conditions, but no peak was detected at that retention time.

The reference solution, however, gave rise to 2 peaks of ~180 milliabsorbance (mAbs) (220 nm) at elution times of 1.545 minutes and 1.737 minutes (35 mAbs [254 nm]) at 1.55 minutes and 1.737 minutes, and 28 mAbs [274 nm] at 1.544 minutes and 1.737 minutes; figure not shown). Comparable peak heights and peak areas were observed for the Codan samples, whereas for 1 Codan sample, significantly higher values for peak heights (244 mAbs at 220 nm, 48 mAbs at 254 nm, and 38 mAbs at 274 nm) and peak areas were encountered at those retention times (Fig 7).

The absorbance leads to the conclusion that there is a higher concentration in Codan samples.

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Coagulation Profile of the Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient No.</td>
<td>Platelets, × 10^3/mL</td>
</tr>
<tr>
<td>1</td>
<td>451</td>
</tr>
<tr>
<td>2</td>
<td>105</td>
</tr>
<tr>
<td>3</td>
<td>318</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>243</td>
</tr>
</tbody>
</table>

APTT indicates activated partial thromboplastin time; PT, prothrombin time; ATIII, antithrombin III; APC, activated protein C; NA, not available.

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Solutions Administered Before Onset of CR-DVT, Anticoagulant Treatment, and Patient Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient No.</td>
<td>Solutions Administered</td>
</tr>
<tr>
<td>1</td>
<td>TPN, thiopental, dobutamine, dopamine, antibiotics, ranitidine, furosemide, fentanyl, labetalolhydrochloride</td>
</tr>
<tr>
<td>2</td>
<td>TPN, dobutamine, antibiotics, ranitidine, furosemide, sodium valproate, intralipid</td>
</tr>
<tr>
<td>3</td>
<td>TPN, dobutamine, dopamine, sodium valproate, phenobarbital, midasolam, Mor, antibiotics, ranitidine, furosemide, paracetamol</td>
</tr>
<tr>
<td>4</td>
<td>TPN, alizapride, tramadol hydrochloride, antibiotics, ranitidine, furosemide, IL, vitamin K</td>
</tr>
<tr>
<td>5</td>
<td>TPN, ranitidine, dobutamine, antibiotics, midasolam, morphine, cistracurium</td>
</tr>
</tbody>
</table>

TPN indicates total parenteral nutrition; LMWH, low molecular weight heparin.

a Common for all 5 patients.
detector output showed the same spectrum each time as that recorded for the saturated solution of DEHP in water and the reference solution.

The unambiguous identification of TOTM from other organic impurities is difficult, because 2 substances differing in structure can exhibit the same ultraviolet spectrum. Six Codan samples were investigated with HPLC-DAD and LC-MS-DAD. The tubings used to collect the samples were immersed in 200 mL of acetonitrile for 2 hours and placed on the ultrasonic bath. Both Codan and Perfusend tubings consisting of PVC-DEHP were used as softening agents, and both components dissolved in acetonitrile. To obtain better peak shapes, the mobile phase conditions were changed to gradient elution (Table 4).

Injection of the solution of the tubing using HPLC-DAD gave rise to a very large peak at 18.298 minutes elution time. The spectrum (figure not shown) is identical to the saturated solution. Afterward, the Codan samples were injected using the same chromatographic conditions. All samples showed a peak at that retention time, exhibiting the same ultraviolet spectrum as shown in Fig 8.

The Codan fluid samples collected after 24 hours showed the highest absorbances among them, and their spectra were similar to that recorded for the solution of the Codan tubing itself. The 2 filtered (via 0.2-μm inline screen filter) Codan samples collected after 96 hours only showed intermediate peak heights. These observations suggest that the Codan tubing displays an impressive leakage of an organic substance into the infusion solution compared with the Perfusend tubing or reference solution, moderately mitigated by the presence of a submicron inline filter.

To be sure of the nature of the component behind the...
observed retention time, additional structural characterization was necessary. Therefore, the experiments were repeated by using LC-MS-DAD, so that through MS the \( m/z \) value of DEHP could be also determined. DEHP has a molecular weight of 390.6 whereas TOTM exhibits a molecular mass of 546.7 Two stationary phases with similar selectivity were used, applying the same gradient conditions as for the HPLC-DAD experiment.

For the solution of the tubing immersed in acetonitrile injected on the Hypersil C18-BDS, a large peak is shown at retention time 19.49 minutes, of which the ultraviolet spectrum is visualized in Fig 8. The MS scan showed a peak of 100% abundance for the \( m/z \) value of 391, whereas no abundance was measured at \( m/z \) equal to 547 (Fig 9), which strongly suggests the presence of DEHP.

Similarly, the solution of the tubing immersed in acetonitrile injected on the Zorbax SB-C8 stationary phase gave a large peak at retention time 15.68 minutes (Fig 10). Scanning the MS spectrum revealed a peak of 100% abundance at the \( m/z \) value or 391 at that elution time, whereas for the \( m/z \) value of 547, no abundance could be detected (figure not shown).

Minimal Clinical Data Set

For the PICU, the data revealed 3 patients of 117 admissions in 2001 semester 1 and 1 patient of 142 admissions in 2000 semester 1. In the other semesters, no CR-DVT was registered (Fig 8).

The incidence of CR-DVT in the 2001 semester 1 was tested with an exact test for binomial distributions and was significantly high (\( P < .0001 \)).

For the adult ICUs and hospitalization wards for adults, a peak in CR-DVT could be observed in 2001 semester 2 (Fig 9). A total of 52 CR-DVT incidents were counted on 5336 admissions (1.0%) against an average of 0.6% in the other semesters between 2000 and 2004. This is 16 events more than expected (\( P = .002 \)).

DISCUSSION

During a 33-day period in 2001, we had 5 cases of CR-DVT in children admitted to the PICU. We were unable to identify a pathophysiological cause. We could not implicate the femoral catheters because these have been in use and unchanged in composition for years. None of the patients had evidence for hereditary or acquired hypercoagulability.

These thromboses occurred during a period in which the PICU (following hospital budgetary politics and recommendation) changed to using new types of intravenous administration sets and intravenous volumetric PF pumps. This change occurred gradually with a blending of both the different types of PF pumps (Terumo and Argus) and the intravenous tubing sets (Codan, Terumo, and Perfusend). The Terumo PF pumps were progressively replaced with the newer type of Argus PF pumps, while the Codan tubings had invaded local unit stocks hospitalwide, gradually becoming prominently available as the hybrid alternative to the more expensive Terumo tubing. As Argus PF pumps (Perfusend tubing) progressively became more available, there was no additional need to use the Terumo hybrid tubing (Codan) that by now had completely replaced the Terumo tubing. From a retrospective investigator’s point of view, an insuperable shortcoming that inhibits all attempts to statistical analysis is that the type of PF pump and the type of
intravenous tubing used for the administration of intra-
venous fluids were never recorded in the patient or
nursing charts. In fact, an amalgam of 2 different tubings
combined with 2 different types of PF pumps and even
more types of syringe pumps were simultaneously being
used bedside at the time of the occurrence of the 5
CR-DVT incidents. In a best-case scenario, these 5 pa-
tients at the PICU may well have never been infused
over Codan tubings and suffered CR-DVT purely because
of coincidence.

The results of this in vitro experiment, however in-
disputable, establish that the Codan tubing disintegrated
during normal use with a PF pump. This now explains
why bedside nurses throughout the hospital occasionally
observed and reported that Terumo PF pumps did not
administer intravenous fluids as accurately as in the past,
whereas service engineers were unable to detect mal-
functioning of the equipment. SEM documented that
the disintegration of the Codan tubing’s pump segment
occurred as soon as after 6 hours of running time. The in
vitro HPLC experiment determined that the disintegra-
tion product in the Codan PVC tubing is DEHP. The loss
of PVC wall integrity and shedding of plastic into passing
fluid undoubtedly resulted in altered diameters of the
calibrated Codan tubing’s pump segment. As the Codan
tubing flattened and the inner diameter increased, more
fluid was pumped at the initial volume set, causing the
PF pumps to run inadequately.

The incentive to perform this investigation followed
the highly unusual occurrence of CR-DVT in 4 children
in such short period. This led to the assumption that
disintegration of intravenous tubing into intravenous
fluid resulted in actual intravenous administration of
plasticized debris to patients. A fifth patient was reported

FIGURE 8
Ultraviolet spectrum recorded at 19.49 minutes elution
time on the Hypersil C18-BDS using gradient elution con-
ditions for the injection of the solution of the tubing in
acetonitrile.

FIGURE 9
MS spectrum with 100% abundance for m/z value of 391
(identifying DEHP) on the Hypersil C18-BDS using gradi-
ent elution conditions for the injection of the solution of
the tubing in acetonitrile.
soon after this hypothesis was formed but before inline filters were introduced systematically. The in vitro experiment confirmed that the particles shed from the intravenous tubing into the intravenous fluid are of such large size (34.5 ± 6.1 μm) that hypothetically they can easily induce mechanical thrombosis. Moreover, the absence of additional occurrences of CR-DVT at the PICU because of the introduction of 0.2-μm inline filters suggests that the intravenous administration of large plastic particles and DEHP might actually have played a pathophysiological role in the formation of CR-DVT. It is also probable that phthalate plasticized tubing with a higher shore and better elasticity characteristics is likely to release fewer particles. The presence of very large particles on SEM, migrating freely on the surface of the intravenous Codan tubing with disintegration of the inner wall, while other types of investigated tubings (Terumo, Perfusend, and Cardinalhealth) kept their wall integrity for a longer period than 96 hours (Figs 3–6), enforces this hypothesis. HPLC documented that the in vitro insertion of a commercial 0.2-μm inline filter mitigates the phthalate dose to intermediate values, which also may explain the absence of CR-DVT after the filters were introduced in vivo.

Whether or not these 5 patients were infused with considerable amounts of phthalate and PVC particles cannot be confirmed. The administration of medication and fluids via different types of PF and syringe pumps or different types of tubing has never been systematically registered in the PICU nursing plans or in the PICU medical files. More surprising, however, is the observation that during the period 1997 to 2002, Codan tubings were used extensively throughout the hospital. So-called r41 tubings of Baxter without pump segment were used for off-pump gravitational infusion therapy, which implies that almost all Codan tubings were used in the devastating combination with a PF pump. Through the hospital and, therefore, including the PICU, an unrecorded percentage of patients have been infused with medication or intravenous solutions via 44 568 Codan tubings and PF pumps and, therefore, might have been infused with DEHP plasticized PVC particles. According to our data, the Codan tubing is a very morbid one but the most important shifts in the intravenous tubing field occurred in 1998; however, the 5 CR-DVT cases at the PICU were observed in January 2001, which is ~2 years out of date (Fig 11). However, analyzing the hospital Minimal Clinical Data Set (coded with International Classification of Diseases, Ninth Revision, Clinical Modification) for total CR-DVT occurrence shows a surpris-
ingly high incidence in 2001 (semester 2 \( n = 52 \) > semester 1 \( n = 40 \)) compared with the expected \( n = 36 \) on a yearly basis. The short delay in peak of CR-DVT is probably because of the fact that the Codan tubings remained at use at the adult wards until exemption of stock, whereas the PICU systematically introduced inline filters combined with an early and complete ban of the Codan tubings (Fig 12).

The percentage of mechanically pumped intravenous therapies has gradually increased over the years. The hospital material databases (hardware, orders, and medical disposables) reveal how the number of PF and syringe pumps increased and subsequently how the use of intravenous disposables proliferated, while the turnover of patients is still increasing. The overall intravenous tubing order increased 41% over the past 9 years (114 180 in 1996 vs 193 404 in 2005). The hospital as a whole has become more intravenous technical and likely more prone to intravenous technical-related adverse effects. Although this might partially explain the sudden increase of hospital CR-DVT in 2001, it does not explain why until today, the incidence fell back to the expected baseline coincidentally with the Codan tubings and Terumo PF pumps disappearing from the hospital scene. All these findings suggest that the problems started either as the Terumo tubing completely was replaced or when the Codan tubing for unidentified reasons started to disintegrate right before 2001. When the SEM results became available, the Terumo PF pumps and Codan tubings were immediately removed from the PICU, and the hospital management was informed of the preliminary results.

Our in vitro experiments strongly suggest that a considerable number of patients, both at the PICU and throughout the hospital, have been intravenously injected with DEHP plasticized particles. This occurred involuntarily because of wear of the Codan PVC tubing and happened insidiously. A major concern is whether and how this jeopardized the health of the unidentified patient population of that time.

DEHP is the only plasticizer approved by the US Food and Drug administration for medical use.\(^a\) It is a colorless oil used in high concentrations (20%–50%) in medical disposables to soften a commercial resin in white powder form that is called PVC.\(^b\) Despite this approval, serious concerns have been postulated over the potential of adverse effects to health when individuals are exposed to phthalates,\(^16\)–\(^24\) hence the ban on some specific phthalates in children’s toys instituted by the European Commission in June 2005 and fully supported by the European Parliament.\(^25\) To be more specific, 6 of the currently used phthalates to soften PVC toys and childcare articles are now permanently banned for reasons of adverse health effects and reprotoxicity. Under the directive, 3 of these phthalates are classified as undoubtedly toxic: DEHP, dibutyl phthalate, and butyl benzyl phthalate. As for the other 3 phthalates, di-isononyl phthalate, di-isodecyl phthalate, and di-n-octyl phthalate, there is uncertainty of the risks they present. Regardless of age, children should no longer be exposed to toys or childcare articles softened with 1 of these 6 plasticizers (in a concentration >0.1%). From now on, there is a stable legal situation (2005/84/EEC) that will enable industry to plan in conditions of certainty. The earliest possible deadline by which the companies will be required to comply with these new restrictions is likely to be October 2006.\(^26\)

Because phthalates are not covalently bound, they passively leak from the plastic matrix influenced by factors like temperature, exercised pressure, and storage time.\(^27\)–\(^34\) This leakage is influenced by contact with fluids and by flow rates, and there is documented interaction with body fluids like blood, mucus, and saliva,\(^14\)–\(^19,22,25\) hence, the ban for DEHP plasticized toys that can be chewed on and put in a child’s mouth. Remarkably however, and despite this evidence, there is no such ban for the use of DEHP-softened PVC medical devices, which are also put into children’s mouths or even beyond.\(^15,18,22\) The PF pump mechanism is a chewing one because it continuously squeezes the softened PVC tubing between its peristaltic moving fingers and a solid metal plate. Blood as a body fluid is known to interact strongly with spallated or shed DEHP and particularly its monooester metabolites.\(^6,13\) Protein adsorption occurs within seconds of blood–material contact and plays an important role in the subsequent blood response. Increased fibrinogen adsorption can be regarded as the initial index of blood reactions and leads to higher thrombogenicity. This also reflects in higher thrombin-antithrombin III complex and complement C3a values.\(^28\)

This overall relationship among blood and biomaterials may explain the sudden occurrence of catheter-related DVT in the 5 patients of the PICU or the sharp increase of CR-DVT incidence in the hospital in 2001. Furthermore, DEHP and its metabolites are known endocrine disruptors modifying carcinogenesis in rats and mice\(^19\)–\(^24,36,37\) and are genotoxic to human lymphocytes and human mucosal cells.\(^38\) DEHP and its metabolites are also known xenobiotics mimicking or antagonizing sex hormones, hence their established reprotoxicity.\(^19,39\)–\(^44\) It is also remarkable that since plasticized products became ubiquitous in the developed world, asthma and allergies evolved to the status of major health care problems. Evidence now reveals that, besides lifestyle and demographic factors, asthma and certain allergies can reflect a biological response to phthalates and especially DEHP.\(^20\)–\(^23\)

The clinical International Classification of Diseases coded hospital data recorded the unusual high incidence of CR-DVT quite simultaneously with their real-time occurrence. In the PICU, however, there was a documented underscore compared with the real frequency of
CR-DVT. The accurate and a more systematic use of this data set for the monitoring of sentinel events should be considered to reveal problems in patient safety at the earliest moment.45–52 Sentinel function of whatever tool based on entered diagnosis implies that the diagnosis is actually made and put in the registration system. Occult occurrences of conditions or missed diagnosis will not add to any database with new input. In our observations, the oldest (15 years old) of the 5 patients with CR-DVT remained asymptomatic and the thrombosis was coincidentally documented via a computed tomography scan of the pelvic region for another indication (acute abdominal pain). With underdiagnosed CR-DVT, the sentinel function will have its limitations and the true extent of the medical problem remains uncertain.

The in vitro experiments set up retrospectively to investigate the sudden occurrence of 5 CR-DVT incidents at the PICU should have been performed before the Codan tubing was introduced for medical use. The prospective in vitro protocols run by the hospital engineering department focused on PF pump accuracy and may be too scanty. When nurses start to report about PF pump performance inaccuracy, experiments such as those described previously should be initiated to determine the exact cause for this observation. After all, we are dealing with fluids that will be administered intravenously over medical devices softened with hazardous plasticizers.

LIMITATIONS

Because of limited financial resources, the sample size was small. Additional investigation via Fourier transform infrared microscopic identification or energy dispersive radiograph analysis could not be budgeted. No screening for immunoglobulin E anti-DEHP was performed in patients.

To meet with these restrictions, all tubings were collected randomly from different sites in the hospital. All PF pumps involved in these experiments were collected randomly from different sites in the hospital and the serial numbers were registered.

ACKNOWLEDGMENTS

This study was approved by the Bioethical Committee of the Free University of Brussels, Belgium; Ethical Committee ref 2005/157.

We thank K. Gillisjans (Free University Brussels) and W. Janssens (Johnson and Johnson Pharmaceutical Research and Development, a Division of Janssen Pharmaceutica N.V.) for their excellent technical assistance. The technical assistance by the group of E. Maynard and colleagues (Pall Industries, Portsmouth, United Kingdom) is highly appreciated, as are the technical reviews by P. Bauwin (Belgian Federal Government: Public Health, Medical Drugs/Medical Devices Department).

REFERENCES

22. Shea K. American Academy of Pediatrics Committee on Envi-


Childhood *Helicobacter pylori* Infection and Growth Impairment in Developing Countries: A Vicious Cycle?

Henry J. Windle, PhD\textsuperscript{a}, Dermot Kelleher, MD, FRCP\textsuperscript{a}, Jean E. Crabtree, DPhil, MRCPath\textsuperscript{b}

\textsuperscript{a}School of Medicine, Trinity College Dublin and Dublin Molecular Medicine Centre, Dublin, Ireland; \textsuperscript{b}Leeds Institute of Molecular Medicine Unit, St James’s University Hospital, Leeds, United Kingdom

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

**ABSTRACT**

We hypothesize that infection with the gastric pathogen *Helicobacter pylori* in children in developing countries is the initiator of a vicious cycle of events that result ultimately in malnutrition and growth impairment. Acute infection with *H pylori* is accompanied by hypochlorhydria, which facilitates the acquisition of other enteropathogens because of removal of the gastric acid barrier, which then results in diarrheal disease and iron-deficiency anemia. This is likely to occur most frequently in developing regions where the prevalence of *H pylori* infection is disproportionately high and multiple enteric coinfections are common. The consequent synergistic impact of diarrheal disease and micronutrient deficiency on growth and cognitive function in children has significant public health implications for socioeconomic development in these countries.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2196
doi:10.1542/peds.2006-2196

**Key Words**

*H pylori*, hypochlorhydria, diarrhea, enteric infections, iron-deficiency anemia, malnutrition

**Abbreviations**

IDA—iron-deficiency anemia

IL-1—interleukin 1

Accepted for publication Oct 3, 2006

Address correspondence to Jean E. Crabtree, DPhil, MRCPath, Leeds Institute of Molecular Medicine, Level 9 Molecular Medicine Section, St James’s University Hospital, Leeds LS9 7TF, United Kingdom. E-mail: j.crabtree@leeds.ac.uk

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
The acquisition of Helicobacter pylori infection occurs primarily in early childhood. Acute infection, at least in adults, results in a transient or extended period of hypochlorhydria. We hypothesize that this period represents a critical phase for pediatric cohorts in developing countries because, within this window, the individual becomes more susceptible to (1) acquiring other enteropathogenic infections, which results in diarrheal illness, and (2) developing iron-deficiency anemia (IDA). The combined impact of these extragastric manifestations of H pylori infection ultimately results in impaired childhood growth and cognitive function caused by the comorbidity associated with malnutrition, micronutrient deficiency, and diarrheal disease.

Evidence for Our Hypothesis

Since the discovery of H pylori by recent Nobel Laureates Marshall and Warren in 1982, the pathogen has been shown to be a causative agent of disease states of varying degrees of severity including chronic gastritis, peptic ulcer disease, gastric adenocarcinoma, and gastric mucosa-associated lymphoid tissue (MALT) lymphoma. Host genetics, host immune responses, and bacterial virulence factors contribute to the multifactorial nature of disease progression. Interleukin (IL)-1β is a potent suppressor of gastric acid secretion, and polymorphisms in the IL-1 gene cluster resulting in high IL-1β production are associated with hypochlorhydria with chronic H pylori infection. More severe disease is also associated with specific bacterial factors such as the cag pathogenicity island and carriage of specific alleles of the vacuolating cytotoxin.

Although the discovery of H pylori heralded a major reevaluation of the etiology of peptic ulcer disease and distal gastric cancer, the major global impact of H pylori infection may yet be largely unrecognized. In developing countries the incidence of H pylori infection in infancy is high and has been associated with malnutrition and growth faltering. Indeed, the incidence of H pylori infection in malnourished children is greater than in adequately nourished subjects. Moreover, there is an inverse relationship between early H pylori colonization, infant malnutrition, and socioeconomic status. It is believed that the acquisition of H pylori in infancy/childhood in developing countries has a more severe impact on general health compared with developed countries. We believe that the acquisition of H pylori in childhood in developing countries is the initiator of a vicious cycle of events that makes an impact on childhood morbidity and mortality, as summarized in Fig 1.

First, the association between acute H pylori infection and transient hypochlorhydria in adults is well documented. Hypochlorhydria usually resolves within several months. Experimental infection of animals with gastric Helicobacter species also results in initial hypochlorhydria, which in gerbils was reversible by treatment with recombinant IL-1 receptor antagonist. It is interesting to note that animal models have shown that hypochlorhydria can promote fecal transmission of infection, an observation that may have implications for H pylori transmission among humans. However, although there is less direct evidence to support an association between acute H pylori infection and hypochlorhydria in pediatric cohorts, accumulating evidence indicates that H pylori–infected children also have impaired gastric acid secretion.

Second, hypochlorhydria increases susceptibility to enteric infections such as typhoid and nontyphoidal salmonellosis, cholera, giardiasis, and other infections. Other organisms (eg, Escherichia coli, Shigella flexneri) are less susceptible to gastric acid and can survive exposure to acidic pH. Evidence from case-controlled studies that examined the association between proton pump inhibitor–induced hypochlorhydria and increased relative risk of acquiring enteric infections including Clostridium difficile, Giardia, and Salmonella supports the view that a reduction in gastric acid secretion may result in the acquisition of infections, as reviewed elsewhere. (see Table 1). Thus, H pylori–induced hypochlorhydria may predispose to enteric infections, particularly in regions of the developing world where enteric infections are endemic. Evidence from develop-

![FIGURE 1](image)

**FIGURE 1**

Potential cycle induced by childhood H pylori infection in developing countries. H pylori infection and hypochlorhydria result in development of IDA and acquisition of other enteric infections, which promotes a vicious cycle of malnutrition and growth impairment. Polymorphisms in the IL-1β gene cluster may control the extent and duration of hypochlorhydria with initial H pylori infection.

<table>
<thead>
<tr>
<th>Infecting Organism</th>
<th>Odds Ratio</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>C difficile</td>
<td>1.9–2.9</td>
<td>47–52</td>
</tr>
<tr>
<td>Gastroenteritis (ND)</td>
<td>1.8–5</td>
<td>53</td>
</tr>
<tr>
<td>Salmonella</td>
<td>4.6</td>
<td>54</td>
</tr>
<tr>
<td>Campylobacter</td>
<td>11.7</td>
<td>55</td>
</tr>
<tr>
<td>Giardia</td>
<td>—</td>
<td>52 and 56</td>
</tr>
<tr>
<td>Gastroenteritis (ND)</td>
<td>3.58</td>
<td>57</td>
</tr>
</tbody>
</table>

ND indicates no infecting organism determined; —, not reported.

---

PEDIATRICS Volume 119, Number 3, March 2007 e755
ing countries indicates that H. pylori infection may increase the risk of acquisition of coinfection with other pathogens including Vibrio cholerae, Salmonella typhi, Shigella, and various parasites. The resultant diarrhea post–acute H. pylori infection leads to malnutrition and growth retardation in children. The recent linking of enteric protozoal infections in developing countries with impaired cognitive function in children further emphasizes the importance of the vicious cycle initiated by H. pylori on childhood morbidity and potential for socioeconomic impairment.

Another potential consequence of hypochlorhydria is IDA. IDA is a major public health issue, particularly in children, in developing countries where the prevalence in these regions exceeds 50%. In addition to poor nutritional sources of iron in many regions, absorption of bioavailable iron can be compromised if there are alterations to gastric acid secretion (eg, see ref 58). Acidic pH is essential for the efficient reduction and complexation of iron (and other micronutrients) before absorption.

A possible causal role for H. pylori in the onset of IDA is supported by epidemiologic studies in adults and children. A combination of the (1) hypochlorhydria–associated H. pylori infection and (2) direct competition between H. pylori and the host for iron are the likely main contributors to IDA. In developing regions, the micronutrient deficiency and clinical sequelae would be further exacerbated by malnutrition and concurrent enteric infections. Manifestations of clinically advanced IDA include increased childhood mortality and susceptibility to disease, reduced growth, and cognitive function. However, the impact of H. pylori on acquisition of diarrheal pathogens and pediatric growth in developing countries has not yet been evaluated rigorously. Studies in developed countries on the role of H. pylori in diarrheal illness have yielded conflicting results compared with those in developing countries. This is perhaps not surprising, given the markedly contrasting exposure to diarrheal pathogens in developing countries. To date, few studies have focused on children during the first few years of life, which is when acute H. pylori infection and associated hypochlorhydria occur.

**CONCLUSIONS**

A combination of micronutrient deficiency and coinfection with diarrhea–inducing enteropathogens acquired as a consequence of H. pylori–induced hypochlorhydria is likely to have a profound impact on pediatric populations in developing countries where the prevalence of H. pylori is high and reliable nutritional sources of bioavailable iron are low. Thus, prevention of H. pylori infection could potentially have an important impact on diarrheal diseases in the developing world.

**TESTING THE HYPOTHESIS**

Whether the hypochlorhydria induced by acute H. pylori infection in childhood is associated with an increase in enteric infections and diarrheal disease, IDA, and growth impairment remains an important unanswered question. The impact of IL-18 gene cluster polymorphisms on acute H. pylori infection–related hypochlorhydria in pediatric populations is unknown. These factors require evaluation by a multicountry longitudinal observational prospective cohort study in developing countries to determine the impact of H. pylori and other enteropathogen infections on the epidemiology of diarrheal disease, IDA, and childhood growth. As a causative agent of human disease, the global impact of H. pylori on childhood morbidity in developing countries may far outweigh its role in causing peptic ulcer disease and gastric cancer.

**ACKNOWLEDGMENTS**

Work in the laboratories of Drs Windle and Kelleher is supported by the Health Research Board, Enterprise Ireland, the European Union, and the Higher Education Authority. Work in the laboratory of Dr Crabtree is supported by Yorkshire Cancer Research and the European Union. We are funded under the Sixth Framework Program of the European Union, Project CONTENT (INCO-CT-2006-032136).

**REFERENCES**


71. Oppenheimer SJ. Iron and its relation to immunity and infec-
Recurrent Expressive Aphasia as a Presentation of Cat-Scratch Encephalopathy

James W. Fox, MDa, Joanna K. Studley, MDb, Daniel M. Cohen, MDc

aDepartment of Pediatrics, Division of Emergency Medicine, Children’s Hospital of Akron, Akron, Ohio; bDepartment of Pediatrics, cDivision of Emergency Medicine, Columbus Children’s Hospital, Columbus, Ohio

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

ABSTRACT
Cat-scratch disease is a common disease, occurring in an estimated 24,000 patients annually in the United States, and is one of the most common causes of chronic lymphadenitis in children. A wide array of neurologic complications occurs as a result of cat-scratch disease. However, there have been no reports of acute-onset, self-resolving, recurrent, expressive aphasia, as we report here in an adolescent boy. In our case, establishing the diagnosis of cat-scratch encephalopathy saved time and resources and afforded the family a benign diagnosis. Cat-scratch encephalopathy must be considered in the differential diagnoses when pediatric patients present with unusual neurologic symptoms.

FIRST RECOGNIZED BY Foshay in 19321 and reported by Debre et al2 in 1950, cat-scratch disease (CSD) is an infection that typically results from introduction of a pleomorphic Gram-negative bacillus, Bartonella henselae, into the bloodstream via the scratch or bite of an infected cat, usually of young age. The infection most often occurs in the pediatric age group; affected individuals exhibit low-grade fever, regional lymphadenopathy, and malaise. Atypical cases of CSD, however, may affect the central nervous system (CNS). This rare presentation (0.17%–2% of cases4) usually manifests as an encephalopathy with sudden-onset seizures or altered consciousness. It is interesting to note that since the first description of CNS involvement in CSD by Stevens5 in 1952, there have been numerous reports in the literature that described a potpourri of neurologic sequelae associated with CSD, including cerebral arteritis6 and Brown-Sequard syndrome.7 Recently, we cared for an adolescent boy whose case of cat-scratch encephalopathy adds to the intriguing list of neurologic presentations of CSD. To our knowledge, an intermittent, expressive aphasia has never been reported in a patient with cat-scratch encephalopathy. In this report, we describe our patient’s presenting symptoms, the differential diagnoses considered, and his clinical course.

CASE REPORT
A 15-year-old boy with no significant past medical history was evaluated initially at a local community hospital for the complaint of “inability to speak.” Earlier in the day, this straight-A student was sent home from school because of nausea and vomiting. Once at home, his parents noticed that he was unable to communicate verbally. He never experienced typical seizure-like activity, became confused, lost consciousness, lost his ability to understand others, or lost the ability to write. At the community hospital, his examination revealed a well-appearing adolescent with an expressive aphasia, mild right facial droop, right-sided weakness (upper greater than lower extremity), and an inability to ambulate. A complete blood count, electrolytes, renal function tests, and computed tomography of his head were unrevealing. As a result, the referring hospital performed magnetic resonance imaging (MRI) of his brain, the results of which were abnormal (Fig 1; initial interpretation: “diffuse white matter periventricular signal abnormality”). A cerebrovascular insult was the physicians’ foremost concern, so aspirin therapy was begun, and he was admitted overnight. Over the course of 6 hours, his symptoms steadily improved and, by morning, had com-
pletely resolved. He was discharged later that day on aspirin therapy, with a scheduled follow-up appointment by a pediatric neurologist.

Within hours of discharge, the patient had a recurrence of his inability to speak. Subsequent to representing to the referring hospital, he was transferred to the emergency department (ED) of our children’s hospital. Before transfer, intravenous lorazepam was given for “agitation” without noticeable improvement. Our physical examination confirmed findings from the previous day. Specifically, he was a pleasant, cooperative, well-appearing young adult with a pronounced expressive aphasia, right central seventh cranial nerve paresis, right-sided weakness (upper greater than lower extremity), unsteady gait secondary to weakness without cerebellar signs, normal reflexes, no clonus, and extensor Babinski responses bilaterally. In addition, a large (5 × 3 cm), mildly tender, nonerythematous, nonsuppurative, right axillary lymph node was palpated. A few healing papules and scratches on his chest were also identified. This led to the discovery of the patient’s frequent interactions with his family’s kitten, which often scratched him.

Review by a pediatric neuroradiologist of the brain MRI performed by the referring facility revealed the following findings: “symmetric abnormal increased T2-weighted signal and decreased diffusion within the white matter of the centrum semiovale bilaterally including involvement of the subcortical U-fibers.” This interpretation made ischemic injury highly unlikely. While in the ED, the patient displayed a remarkable steady recovery on reevaluation. Nevertheless, given the recurrent nature of his impressive symptoms, he was admitted to our hospital for additional evaluation and observation.

An extensive inpatient evaluation followed. His workup was negative for metabolic disorders, hypercoagulability, thromboembolism, toxic ingestions, toxin exposures, autoimmune disorders, encephalitis, and evidence of traumatic injury. A pediatric psychiatrist opined that his symptoms were not the result of mental illness. During the boy’s hospitalization, the inpatient service continued the aspirin started by the referring institution. In addition, intravenous steroids were administered because the patient’s brain MRI suggested an acute disseminated encephalomyelitis. The following morning, results of the CSD serum titers sent from the ED were strongly positive (IgM > 1:32 and IgG > 1:128).

In light of the otherwise negative workup, this finding confirmed the diagnosis of CSD encephalopathy. Per our infectious disease consultant, a 5-day course of azithromycin was instituted.

Within hours of hospitalization and recrudescence of his symptoms, the patient made a second complete recovery. He has remained symptom free, without recurrence or need for follow-up with a pediatric neurologist, 5 months after his hospitalization.

FIGURE 1
Brain MRI in a patient with cat-scratch encephalopathy. Significant diffuse inflammation in the subcortical white matter is evident in these T2-weighted images, which were obtained during our patient’s initial presentation.
DISCUSSION

When considering the cause of expressive aphasia in children, physicians commonly consider cerebrovascular ischemia or infarction, partial seizures, cerebral malignancies, traumatic injuries, complex migraines, vasculitic syndromes, and psychiatric illnesses. We have expanded this differential diagnosis by describing a child with cat-scratch encephalopathy who presented primarily with expressive aphasia.

In general, CSD is a self-limited, subacute regional lymphadenitis caused by B henselae. Atypical presentations occur in 11% of cases, with encephalopathy among the most common. Encephalopathy accounts for 90% of CSD involvement of the CNS and typically manifests as seizure activity (frequently difficult-to-control status epilepticus), headache, coma, and combative behavior. Other reported CNS complications include cerebral arteritis, peripheral and central facial nerve paralysis, Brown-Séquard syndrome, neuroretinitis, hemiparesis, cerebellar ataxia, and movement disorders. It is notable that results of polymerase chain reaction for B henselae were negative in our patient. This is, fortunately, a rare occurrence.

The mechanism by which cat-scratch encephalopathy occurs is not known. Previous authors have suggested direct invasion of the CNS, release of neurotoxin, and an immune-mediated vasculitis. It is surprising that although a bacterium is the etiologic agent of this encephalopathy, the seizures and coma that typically develop occur this is, fortunately, a rare occurrence.

Cat-scratch encephalopathy rarely presents without a change in mental status (seizure, coma, combative behavior), and aphasia as a symptom of cat-scratch encephalopathy has only been reported twice before in the literature. The first case involved a 25-year-old woman with CSD who was hospitalized after a seizure. After 4 days of treatment, she became lucid enough to speak and, at that time, demonstrated both a receptive and expressive aphasia. In contrast to our patient, aphasia was not her presenting symptom. The other case involved a 7-year-old girl who presented with an expressive aphasia and acute right hemiplegia, similar to our case. This patient demonstrated a protracted course, as her aphasia lasted nearly 2 months and her right hemiplegia was still present after 5 months of follow-up. Therefore, ours is the first report of cat-scratch encephalopathy presenting with an expressive aphasia of very short duration (hours).

Another rare occurrence in our patient was the recurrent nature of his symptoms. Only twice before in the literature was a recurrent encephalopathy attributable to CSD described. One patient was a 14-year-old girl who experienced 2 qualitatively different seizures 3 weeks apart. In addition, she continued to suffer from persistent speech and language difficulty 10 weeks after discharge. The other case that demonstrated an element of recurrence involved a 17-year-old girl who was admitted twice, 4 days apart, for generalized seizures that were qualitatively the same. She experienced complete resolution of her symptoms between convulsions and, again, after her second seizure. Our patient represents the first report of a case of CSD with recurrence of the same neurologic abnormality separated by complete resolution within a 24-hour time frame.

It is possible that our patient’s neurologic symptoms were a manifestation of nonconvulsive partial status epilepticus. There are several reports in the literature of aphasia representing the predominant symptom (often with right-sided motor deficits, as in our patient) of nonconvulsive partial status epilepticus. It is interesting to note that partial seizures remain a potential cause of our patient’s aphasia despite his lack of response to intravenous lorazepam. In the setting of cat-scratch encephalopathy, seizures are notoriously difficult to control with antiepileptic medications. Furthermore, reports on aphasic status epilepticus have demonstrated this same difficulty. Abnormalities shown by MRI, electroencephalography, and single-photon emission computed tomography typically localize to the left frontotemporal region in patients with aphasic status epilepticus. Although it is difficult to localize the focus of our patient’s symptoms from a neuroanatomic perspective because of the diffuse, bilateral abnormalities seen on his MRI, the suspect frontotemporal region was certainly involved.

It is notable that results of polymerase chain reaction testing of the cerebrospinal fluid for B henselae were negative in our patient. This is consistent with previous reports of patients with cat-scratch encephalopathy whose cerebrospinal fluid tested negative for B henselae by polymerase chain reaction.

During his hospitalization, the patient was treated with intravenous corticosteroids because his brain MRI suggested diffuse cerebral inflammation. Although it is unlikely that this treatment resulted in the prompt disappearance of his symptoms, literature does exist that supports the use of corticosteroids for cat-scratch encephalopathy. In addition, his therapy included a course of azithromycin. Although the efficacy of antimicrobial therapy for CSD, and even more so for cat-scratch encephalopathy, is controversial, in light of this atypical presentation and the presence of large lymphadenitis, the use of azithromycin was felt to be of potential benefit and of limited to no harm for this patient.

CONCLUSIONS

This case adds to the expanding literature on atypical neurologic complications of CSD. Our report confirms the prediction made by Marra in 1995: “Recent ad-
Advances in identification of *B henselae* will lead to recognition of more neurologic complications.” It is important to note that this case emphasizes the need to inquire about cat exposure when children present with unusual neurologic symptoms. Furthermore, when a clinician is confronted with a child with expressive aphasia, cat-scratch encephalopathy should be considered.

REFERENCES

Marked Hyperbilirubinemia Associated With the Heme Oxygenase-1 Gene Promoter Microsatellite Polymorphism in a Boy With Autoimmune Hemolytic Anemia

Stephan Immenschuh, MDa, Ying Shan, MD, MPHb, Hartmut Kroll, MDa, Sentot Santoso, PhDa, Wilhelm Wössmann, MDc, Gregor Bein, MDb, Herbert L. Bonkovsky, MDk

aInstitute of Clinical Immunology and Transfusion Medicine and bDepartment of Pediatric Hematology and Oncology, Justus Liebig University, Giessen, Germany; cDepartment of Medicine and the Liver-Biliary-Pancreatic Center, University of Connecticut Health Center, Farmington, Connecticut

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

ABSTRACT

Mild hyperbilirubinemia is a clinical feature of hemolysis. Here we describe a boy with marked elevation of serum bilirubin values (maximum: 70 mg/dL) during an acute episode of autoimmune hemolytic anemia, which returned to within the reference range after clinical improvement. The boy was a homozygous carrier of short alleles of the heme oxygenase-1 (HO-1) gene GT dinucleotide-repeat promoter polymorphism, which is associated with increased activity and inducibility of the heme-degrading enzyme HO-1, which catalyzes the production of bilirubin. In addition, heterozygosity of the uridine 5′-diphosphate-glucuronosyl-transferase 1A1 promoter polymorphism that is linked with Gilbert syndrome was found in this patient. Because bilirubin production plays a critical role during the neonatal period, the HO-1 promoter polymorphism may be an important genetic factor for the clinical outcome of neonatal hyperbilirubinemia.

BILIRUBIN is a double-edged sword that, on the one hand, is a potent antioxidant with clinically relevant protective effects but, on the other hand, is potentially toxic. In neonatal hyperbilirubinemia, high concentrations of bilirubin may cause kernicterus and death. Serum levels of bilirubin are determined by a fine-tuned balance of production and elimination and a complex interplay of various genetic factors such as the uridine 5′-diphosphate-glucuronosyl-transferase 1A1 (UGT1A1) TATAA element polymorphism, which regulates the steady-state activity of this key enzyme for bilirubin conjugation. Production of bilirubin is intimately linked to heme oxygenase (HO), which catalyzes the first and rate-controlling step of heme degradation and generates biliverdin that is converted into bilirubin via biliverdin reductase. Genetic deficiency of HO-1, the inducible isoform of HO, leads to anemia, hepatomegaly, and lymph node swelling in knock-out mice, and these clinical features have also been described in a human case of genetic HO-1 deficiency.

Upregulation of HO-1 protects against a variety of pathophysiological conditions such as inflammation and sickle cell crisis. A promoter polymorphism of the human HO-1 gene, which is correlated with and determines the level of expression and inducibility of the endogenous gene, has been shown to be associated with the susceptibility to develop various diseases. Specifically, short alleles of the GT dinucleotide-repeat polymorphism within the promoter of the human HO-1 gene (<25 repeats) cause high levels of HO-1 gene induction in response to oxidative-stress stimuli. In genetic association studies, carriers of short alleles of the HO-1 GT-repeat polymorphism exhibited a lower incidence of smoking-related emphysema, atherosclerosis, and diabetes.

Key Words: bilirubin, oxidative stress, genotype-phenotype correlation, genetic testing, kernicterus

Abbreviations: UGT1A1, uridine 5′-diphosphate-glucuronosyl-transferase 1A1; HO, heme oxygenase; AIHA, autoimmune hemolytic anemia; PCR, polymerase chain reaction

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1385
doi:10.1542/peds.2006-1385

Accepted for publication Sep 27, 2006
Address correspondence to Stephan Immenschuh, MD, Institut für Klinische Immunologie und Transfusionsmedizin, Justus-Liebig-Universität Giessen, Langhansstrasse 7, 35392 Giessen, Germany. E-mail: stephan.immenschuh@immunologie.med.uni-giessen.de

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
tes mellitus. However, little is known regarding the clinical impact of this HO-1 gene polymorphism during hemolysis in which large amounts of heme that are released from erythrocytes may induce HO-1 gene expression. This could be of particular importance in hemolytic conditions such as autoimmune hemolytic anemia (AIHA) and neonatal hyperbilirubinemia.

Here we report a case of marked hyperbilirubinemia during an acute hemolytic episode of AIHA in a boy who is a homozygous carrier for short alleles of the HO-1 gene GT-repeat promoter polymorphism but not for the long TATAA element polymorphism of the UGT1A1 element.

CASE REPORT

A 14-year-old white boy with a history of chronic relapsing autoimmune thrombocytopenia presented with an acute onset of jaundice, pallor, and hepatosplenomegaly. We obtained a blood count, which showed pancytopenia with a hemoglobin level of 3.4 g/dL, white blood cell count of 1.7 × 10⁹/L, and platelet count of 64 × 10⁹/L; the peripheral blood film exhibited spherocytosis and reticulocytosis. Hemolysis was confirmed by additional laboratory findings (initial lactate dehydrogenase: >1007 U/L [reference: <200 U/L]; absent haptoglobin; elevated plasma hemoglobin), and results of direct and indirect Coombs’ tests were positive. No laboratory signs of extrahepatic cholestasis (alkaline phosphatase: 45 U/L [reference: <500 U/L]), liver cell necrosis (alanine aminotransferase: 33 U/L [reference: 10–50 U/L]), or impaired liver function (albumin: 4050 mg/dL [reference: 3500–5200 mg/dL]; pseudocholinesterase: 7054 U/L [reference: 5100–12 600 U/L]) were detected, and the patient was diagnosed with Evans syndrome (AIHA and thrombocytopenia). The boy needed up to 5 packed red blood cell transfusions daily to maintain a hemoglobin level of 4 to 5 g/dL. The initial serum bilirubin level of 18 mg/dL (16.6 mg/dL unconjugated bilirubin) rose to a peak of 70 mg/dL (18 mg/dL unconjugated bilirubin) after 5 days (Fig 1). Total and unconjugated bilirubin levels were measured with an automated system (Architect; Abbott Diagnostics, Wiesbaden, Germany) with commercially available assays (Total Bilirubin [order No. B8G621] and Direct Bilirubin [order No. B8G631]; Abbott Diagnostics), which are based on the diazo method. Treatment with glucocorticoids and intravenous immunoglobulin slightly slowed hemolysis, but transfusion dependency persisted. Two doses of cyclophosphamide had no effect, and rituximab was given 4 times in weekly intervals. After performance of a splenectomy and treatment with cyclosporine, hemolysis improved and serum bilirubin levels of the patient decreased to reference values (0.7 mg/dL), when hemolysis stopped completely. No obvious neuropsychiatric sequelae of kernicterus were observed during the hospital stay of the patient or after discharge.

Genotyping for HO-1 promoter allelic variants was performed with DNA from peripheral blood by polymerase chain reaction (PCR) amplification of the GT repeat containing promoter fragments of the HO-1 gene with forward primer 5’-AGAGCCTGCAGCTTCTCAGA-3’ and reverse primer 5’-ACAAAAAGCTGGCCATAGGAC-3’. After gel separation, PCR-amplified products were sequenced and the numbers of GT repeats were determined. The patient and his brother were homozygous carriers of short alleles (21 and 22 GT repeats; Fig 2). Similarly, the mother was a homozygous carrier (two 22-GT repeats), whereas the father was a heterozygous carrier of short and long alleles (21 and 30 GT repeats). Genotyping for the long UGT1A1 TATAA element mutation, which leads to Gilbert syndrome, with increased serum bilirubin levels was performed with forward primer 5’-GTCACGTCAGACATGAAACATT-3’ and reverse primer 5’-CTCCACAGGCCATGGGCCCTTT-3’ that flank the TA-repeat of the UGT1A1 TATAA ele-

![Figure 1](image1.png)

**FIGURE 1**
Serum bilirubin values in a patient with an acute episode of AIHA. Bilirubin values were determined over 36 days after admission of the patient. Hemoglobin (Hb) values are shown for a comparison.

![Figure 2](image2.png)

**FIGURE 2**
HO-1 promoter allelic variants in the patient and his family. Genomic DNA of the patient and his family was extracted from peripheral blood. Promoter fragments of the HO-1 gene with GT repeats were amplified by PCR with specific primers, and the products were separated on an agarose gel. After extraction of fragments from the gel, the procedure was repeated. Samples were sequenced, and the numbers of GT repeats are indicated by the arrows with the legend below the gel. The patient and his brother were carriers of the 21 and 22 GT repeats, having received one allele from the father (21 GT repeats) and another allele from the mother (22 GT repeats). bp indicates base pairs.
ment. The patient and other family members were heterozygous for the longer UGT1A1 TATAA-box mutation (data not shown).

**DISCUSSION**

We have described a boy with an unusually marked elevation of serum bilirubin values during an acute hemolytic episode of AIHA (Fig 1). The patient was homozygous for the short GT dinucleotide-repeat promoter polymorphism of the HO-1 gene, which is the inducible isoform of the heme-degrading enzyme HO (Fig 2). The GT-repeat polymorphism is associated with an increased expression and inducibility of the human HO-1 gene, which in turn increases the production of bilirubin in a coupled enzyme reaction with biliverdin reductase. It is widely accepted that hemolysis leads to mild hyperbilirubinemia, but serum bilirubin concentrations of >6 mg/dL are not commonly encountered in this clinical condition. Although we cannot rule out the possibility that Evans syndrome (AIHA and thrombocytopenia) may cause more severe hemolysis than AIHA alone, this possibility seems unlikely. Other causes of hyperbilirubinemia such as extrahepatic obstruction, liver cell necrosis, or impaired liver function were not observed in this patient. Thus, to our knowledge, this is the first reported case in which massive hyperbilirubinemia is associated with the short HO-1 gene GT dinucleotide-repeat promoter polymorphism. The “cutoff” length for the short allele of the HO-1 GT-repeat polymorphism is <25 GT repeats, as initially defined by Yamada et al and subsequently confirmed in a number of follow-up studies. Another genetic polymorphism that may lead to hyperbilirubinemia (Gilbert syndrome) is associated with the numbers of TA repeats in the TATA box of the UGT1A1 gene. Genotyping of the UGT1A1 TATAA element in this patient ruled out full-blown Gilbert syndrome, but a potential ancillary effect of limited UGT1A1 activity is not excluded, because the patient was heterozygous for this polymorphism (data not shown). Moreover, because the coding region of the UGT1A1 gene has not been sequenced, compound heterozygosity for UGT1A1 coding and promoter sequence variants cannot be ruled out. Thus, in the present case, homozygosity for the short GT-repeat polymorphism of the HO-1 gene is likely to explain the unusually marked elevation of serum bilirubin levels via an increased supply of the HO-1 inducer heme during an acute episode of AIHA.

Bilirubin protects against various diseases that are related to oxidative stress (for a review see ref 3), and these beneficial effects have been attributed to the potent antioxidant function of this compound. Because HO is the major producer of endogenous bilirubin, the increased inducibility of HO-1 resulting from the short GT-repeat polymorphism may explain the role of HO-1 as a protective genetic factor in such diseases. By contrast, high serum levels of bilirubin can be toxic in neonatal hyperbilirubinemia, in which condition infants with high levels of unconjugated serum bilirubin may develop and die as a result of kernicterus. Therefore, the balance of production and elimination of bilirubin is of particular importance during the neonatal transition period, and an increased inducibility of HO-1 because of the short allele of the HO-1 gene promoter polymorphism may significantly contribute to the severity of neonatal hyperbilirubinemia. A recent study on Japanese infants on this issue did not find an association between the severity of neonatal hyperbilirubinemia and the length of the HO-1 gene GT-repeat promoter polymorphism. The results of the latter report, however, have to be interpreted with caution, because ethnic background seems to be an important factor for the clinical impact of the HO-1 promoter polymorphism.

**CONCLUSIONS**

Unusually marked hyperbilirubinemia during hemolysis was observed in a patient who was homozygous for short alleles of the HO-1 gene GT dinucleotide-repeat promoter polymorphism. Thus, this HO-1 gene polymorphism may be a prognostic factor for the severity of neonatal hyperbilirubinemia and is worthy of additional study in conditions of bilirubin overproduction.

**REFERENCES**

monoxide: from basic science to therapeutic applications. *Physiol Rev.* 2006;86:583–650


Radiologic and Neurophysiologic Aspects of Stroke-like Episodes in Children With Congenital Disorder of Glycosylation Type Ia

Argirios Dinopoulos, MD, PhD, Ismail Mohamed, MD, Blaise Jones, MD, Sanjai Rao, DO, David Franz, MD, Ton de Grauw, MD, PhD

Divisions of Neurology and Radiology, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio

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

ABSTRACT

In an effort to shed light on the mechanism of hemiparetic stroke-like events experienced by patients with congenital disorder of glycosylation type Ia, we evaluated 3 children with this disorder by brain imaging studies and continuous electroencephalogram monitoring during such events. No evidence of ischemia or infarction was revealed on imaging studies and electrographic seizures or intermittent epileptiform activity was demonstrated on electrographic recordings. All 3 patients showed clinical and electrographic improvement after administration of antiepileptic medication. Epileptic phenomena can complicate the stroke-like events of patients with congenital disorder of glycosylation type Ia, and the cause of the hemiparesis may indeed be an active epileptic inhibitory process. As such, electroencephalogram monitoring is warranted, and treatment with anticonvulsant agents is indicated.

Congenital disorders of glycosylation (CDG) comprise a group of inherited multisystem diseases characterized by defects in the glycosylation of glycoproteins. CDG type Ia (phosphomannomutase 2 deficiency) is, by far, the most frequently diagnosed N-glycosylation disease. It has a static clinical course that is characterized by mental retardation, hypotonia, cerebellar dysfunction, polyneuropathy, seizures, and stroke-like episodes of unknown origin. These episodes are known to occur in 40% to 55% of cases and are thought to be triggered by infection. They are associated with irritability, mental status changes, and hemiparesis and are sometimes accompanied by seizures. Their duration varies, with recovery ranging from days to weeks.

In an effort to shed light on the mechanism of these hemiparetic stroke-like events, we evaluated 3 patients with genetically confirmed CDG-Ia by continuous electroencephalogram (EEG) monitoring and brain imaging studies during such events. We report our findings here and propose a possible etiology.

CASE REPORTS

PATIENT 1. A 2½-year-old girl with CDG-Ia presented at the emergency department (ED) 36 hours after acute onset of left hemiparesis, irritability, and a diminished level of consciousness after a febrile illness. At presentation, unresponsiveness and upward eye-rolling were observed. Treatment with intravenous lorazepam (0.2 mg/kg) was followed by respiratory suppression, which required intubation. Continuous bedside EEG monitoring during the hemiparesis revealed right parietal-occipital slowing with frequent intermittent spikes and 6 electrographic seizures arising from the right posterior quadrant. These seizures varied in duration, with the longest lasting up to 15 minutes (Fig 1). With the exception of hemiparesis, there were clinically no motor correlates to these seizures. Imaging studies (Table 1) performed during the hemiparetic event showed no evidence of ischemia or infarction. The patient was treated with intravenous fosphenytoin. Shortly after the initiation of treatment, her EEG showed resolution of electrographic seizures, and her mental status was steadily improved. Hemiparesis completely resolved 96 hours after the onset of symptoms and 24 hours after the administration of intravenous lorazepam.
nous fosphenytoin. The patient was discharged on carbamazepine and has had no additional episodes over the subsequent 4 years.

PATIENT 2. An 11-year-old girl with CDG-Ia presented at the ED 8 hours after acute onset of left-side hemiparesis, irritability, and a diminished level of consciousness concurrent with a viral infection. She had had 3 previous episodes of stroke-like events (at 3, 6, and 9 years of age) that lasted up to 72 hours; because these events occurred in conjunction with seizures, they were described by previous clinicians as Todd’s paralysis. After admission, continuous bedside EEG monitoring showed continuous focal right frontal slowing with intermixed runs of semi-rhythmic spikes and sharp waves that lasted up to 8 seconds (Fig 2). Imaging studies (Table 1) performed during the hemiparesis showed no evidence of ischemia or infarction. Treatment with intravenous lorazepam (0.1 mg/kg) resulted in disappearance of the rhythmic spikes and sharp-wave activity (Fig 2) and complete resolution of the hemiparesis 12 hours later. The patient was discharged on oxcarbazepine and has had no additional hemiparetic episodes over the subsequent 3½ years.

PATIENT 3. A 14-year-old girl with CDG-Ia presented at the ED 48 hours after developing agitation, right-arm weakness, and fever. After becoming increasingly lethargic, she was admitted to the ICU. Bedside EEG monitoring showed continuous rhythmic high-amplitude bifrontal sharply contoured δ activity with higher amplitude on the left frontocentral region. Shortly after the administration of intravenous lorazepam (0.1 mg/kg), rhythmic frontal slowing disappeared and clinical improvement was observed (Fig 3). Hemiparesis resolved 2 hours after the administration of intravenous lorazepam and 72 hours after the onset of symptoms. Computer tomography of the head during the hemiparesis showed no acute...
changes, and both MRI and magnetic resonance angiography performed after the resolution of the hemiparesis showed no evidence of ischemia. The patient was discharged on topiramate and has had no additional hemiparetic episodes over the subsequent 2 1/2 years.

DISCUSSION
Here we report the electrographic and imaging findings of 3 patients with CDG-Ia during a hemiparetic stroke-like event. Signs of neuronal hyperexcitability were demonstrated on the electrographic recordings of all 3 patients, with no evidence of ischemia revealed by the imaging studies. To our knowledge, ours is the first study to document electrographic and imaging findings during hemiparetic stroke-like events.

Several studies have described the neurologic manifestations of CDG-Ia; however, because detailed evaluation during hemiparetic stroke-like episodes is lacking, the pathophysiology and specific mechanism of these episodes remain unknown. Ischemia has been proposed as a possible etiology of these stroke-like events, because both procoagulant (factor XI) and anticoagulant

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Patient Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case No.</td>
<td>Mutations of PMM2 Gene and Coagulation Deficiencies</td>
</tr>
<tr>
<td>1</td>
<td>Mutations: I120T/V231M; coagulation deficiencies: factor XI, antithrombin III, heparin cofactor II</td>
</tr>
<tr>
<td>2</td>
<td>Mutations: R141H/F119L; coagulation deficiencies: factor XI, antithrombin III, protein C, heparin cofactor II</td>
</tr>
<tr>
<td>3</td>
<td>Mutations: R141H/F119L; coagulation deficiencies: factor XI, antithrombin III, protein C and Sc</td>
</tr>
</tbody>
</table>

PMM2 indicates phosphomannomutase 2; HCT, head computer tomography; FLAIR, fluid-attenuated inversion recovery; FSE, fast-spin echo; ADC, apparent diffusion coefficient; HMRS, proton magnetic resonance spectroscopy; MRA, magnetic resonance angiography; MRV, magnetic resonance venography.

a All MRI studies were performed on a 1.5-T magnetic resonance scanner (GE Medical Systems, Milwaukee, WI).
(protein C, antithrombin III, and protein V) proteins are reported to be deficient in most patients with CDG-Ia.4,5 Other factors that support ischemia as the cause of the events include platelet hyperaggregability and transient decrease of endogenous anticoagulants during catabolic stress.6 Nevertheless, there has been no report that provided clear neuroradiologic or pathologic evidence of ischemic stroke, and there have been very few reports of other thrombotic accidents in patients with CDG-Ia.7,8

The above-described coagulation deficiencies were also observed in all 3 of our patients (Table 1); however, no evidence suggestive of ischemia was seen in the neuroimaging studies performed during the hemiparetic events. Moreover, for 2 patients, no restricted diffusion was observed on diffusion-weighted imaging (DWI) or apparent diffusion coefficient maps. Signal-intensity changes can be detected within minutes of arterial occlusion with DWI, and increased signal intensity detected with T2-weighted sequences is usually observed after 8 hours.9 Regardless of the fact that MRI modalities have been shown to miss transient ischemic events, DWI and apparent diffusion coefficient mapping, in particular, usually detect events associated with prolonged duration and disturbances of higher brain function.10 If the prolonged hemiparesis (lasting >48 hours) in our cases was caused by a cortical ischemic infarction, this infarction would have likely been detected on standard MRI. The cytotoxic edema caused by the infarction would be even more likely to exhibit a restricted diffusion on DWI.

Epileptic seizures are not uncommon in patients with CDG-Ia and often are observed in association with stroke-like events. A transient focal neurologic deficit known as postictal paralysis (Todd’s paralysis) can follow a convulsive motor seizure. Our data are insufficient to prove that the paralysis in our patients was ictal; however, the phenomenon cannot be considered a simple Todd’s paralysis, because the hemiparesis persisted beyond 24 hours, ictal discharges were present on EEG, and clinical motor seizures were absent before hemiparesis.11 The EEG recording during the hemiparesis showed clear electrographic seizures in patient 1. No electrographic seizures were captured for patients 2 and 3. However, EEG recording showed a continuous semirhythmic slowing with intermittent runs of spikes and sharp-waves activity in patient 2 (Fig 2) and a continuous sharply contoured rhythm δ activity in patient 3 (Fig 3). Both patterns resolved after the administration of lorazepam. Continuous rhythmic sharp waves and frequent rhythmic spikes and sharp waves are commonly seen in comatose patients but also can be seen in patients with nonconvulsive status epilepticus.12 The significance of such patterns is unclear and reflects the difficulty of differentiating ongoing epileptic discharges from abnormalities caused by diffuse metabolic neuronal dysfunction. The rapid electrographic response and, more importantly, the clinical response observed after treatment with antiepileptic drugs favor an epileptic etiology in our patients.

During the events, electrographic seizures in patient 1 and epileptiform discharges in patients 2 and 3 lateralized with the hemiparesis and, specifically for patients 2 and 3, the maximal activity was observed on the premotor cortex. Inhibitory motor seizures with ictal hemiparesis of variable duration, and even inhibitory motor status epilepticus with hemiparesis lasting from 14 hours to 6 weeks, have been described in the literature.13–15 Penfield and Jasper16 and, more recently, Luders et al17 have reported a negative motor response elicited by stimulation of the human premotor cortex. Despite the inherent limitations of EEG surface recordings, we postulate that the clinical hemiparesis was the result of
subclinical electrographic seizures that actively inhibited motor function.

The pathogenesis of the observed neuronal hyperexcitability in patients with CDG-Ia is unclear, but it is presumably related to a deficient glycosylation at the neuronal level.\(^{18}\) Transient disease-related noncytotoxic edema is also a possibility; Pearl and Krasnewich\(^{2}\) described a patient with CDG-Ia who demonstrated neuroradiologic evidence of transient edema during a stroke-like event. Transient cortical hypoperfusion, induced by the coagulation defects, that was severe enough to cause neuronal dysfunction but mild enough to be undetected by the MRI modalities cannot, however, be excluded. Focal neuronal hyperexcitability has been documented also in patients with mitochondrial encephalopathy lactic acidosis and stroke syndrome and has been proposed as the triggering mechanism of stroke-like events.\(^{19}\)

**CONCLUSIONS**

On the basis of our clinical observations, EEG recordings, and neuroimaging findings, the diffuse and focal neuronal dysfunction that is evident during stroke-like events in patients with CDG-Ia cannot be explained solely by ischemia. Clinical and subclinical seizure activity may complicate the stroke-like events, and the cause of the hemiparesis may indeed be an active epileptic inhibitory process. As such, EEG or, ideally, video-EEG monitoring is warranted, and treatment with anticonvulsant agents is indicated.

**ACKNOWLEDGMENTS**

We thank Dr G. Matthijs from the Center for Human Genetics in Leuven, Belgium, for the mutation screening in the PMM2 gene; Dr M. Privitera for thoughtful comments; and A. P. Cohen, medical writer, for assistance in writing this manuscript.

**REFERENCES**


Citrin Deficiency: A Novel Cause of Failure to Thrive That Responds to a High-Protein, Low-Carbohydrate Diet

David Dimmock, MBBSa,b, Keiko Kobayashi, PhDc, Mikio Iijima, PhDc, Ayako Tabata, MSc, Lee-Jun Wong, PhDa, Takeyori Saheki, MD, PhDc, Brendan Lee, MD, PhDa,b,d, Fernando Scaglia, MDa,b

aDepartment of Molecular and Human Genetics and bHoward Hughes Medical Institute, Baylor College of Medicine, Houston, Texas; cDepartment of Molecular Metabolism and Biochemical Genetics, Kagoshima University Graduate School of Medical and Dental Sciences, Kagoshima, Japan

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

ABSTRACT

The proband was born at 36 weeks, appropriate for gestational age, to nonconsanguineous white parents. There was no evidence of hyperbilirubinemia or intrahepatic cholestasis in the neonatal period, and she had normal newborn screen results. She presented with 3 episodes of life-threatening bleeding and anemia. The diagnostic evaluation for her bleeding diathesis revealed an abnormal clotting profile with no biochemical evidence for hepatocellular damage. She was incidentally noted to have severe growth deceleration that failed to respond to 502 kJ/kg (120 kcal/kg) per day of protein-hydrolyzed formula. An extensive diagnostic workup for failure to thrive, which was otherwise normal, included plasma amino acid analysis that revealed hyperglutaminemia and citrulline levels within the reference range. Testing of a repeat sample revealed isolated hypercitrullinemia. No argininosuccinic acid was detected. Her ammonia level and urine orotic acid were within the reference ranges. Subsequent plasma amino acid analysis exhibited a profile suggestive of neonatal intrahepatic cholestasis caused by citrin deficiency with elevations in citrulline, methionine, and threonine. Western blotting of fibroblasts demonstrated citrin deficiency, and a deletion for exon 3 was found in the patient’s coding DNA of the SLC25A13 gene. On the basis of the experience with adults carrying this condition, the patient was given a high-protein, low-carbohydrate diet. The failure to thrive and bleeding diathesis resolved. When compliance with the dietary prescription was relaxed, growth deceleration was again noted, although significant bleeding did not recur. This is the first report of an infant of Northern European descent with citrin deficiency. The later age at presentation with failure to thrive and bleeding diathesis resolved, when compliance with the dietary prescription was relaxed, growth deceleration was again noted, although significant bleeding did not recur. This is the first report of an infant of Northern European descent with citrin deficiency. The later age at presentation with failure to thrive and bleeding diathesis resolved and without obvious evidence of neonatal intrahepatic cholestasis expands the clinical spectrum of citrin deficiency. This case emphasizes the importance of continued dietary control and growth monitoring in children with neonatal intrahepatic cholestasis caused by citrin deficiency and identifies a new metabolic entity responsible for failure to thrive.

FAILURE TO THRIVE (FTT) is a common pediatric condition, with 10% of children in 1 study showing transitory FTT and 4% showing sustained FTT.1 Investigation of such weight loss is a frequent cause of hospital admissions. Recent reviews have emphasized that this entity most commonly reflects a lack of adequate energy intake and is rarely caused by an organic disease.2 Conversely, citrin deficiency is an inborn error of metabolism previously reported to have a carrier prevalence of 1 in 70 in the Southeast Asian population.3 It has not been described previously in North America or in a subject of Northern European descent. This disorder is caused by mutations in the SLC25A13 gene, which encodes an aspartate glutamate carrier. The same mutations cause 2 different age-dependent clinical phenotypes: neonatal intrahepatic cholestasis caused by citrin deficiency (NICCD [Online Mendelian Inheritance in Man No. 605814])4 and adult-onset type II citrullinemia (Online Mendelian Inheritance in Man No. 603471).5 In both cases, individuals have an abnormal amino acid metabolism.

Key Words: citrullinemia, metabolic disease, liver disease, bleeding disorder, factitious injury

Abbreviations: FTT, failure to thrive; NICCD, neonatal intrahepatic cholestasis caused by citrin deficiency; PT, protime

doi:10.1542/peds.2006-1950

Accepted for publication Sep 18, 2006
Address correspondence to Fernando Scaglia, MD, Clinical Care Center, Suite 1560, 6621 Fannin St, Houston, TX 77030. E-mail: fscaglia@bcm.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
chromatogram with an elevated citrulline level. Severe intrahepatic cholestasis with fatty liver is the most common presenting feature in NICCD. However, neonatal hepatitis; positive newborn screen results for galactosemia, tyrosinemia, and phenylketonuria; hemolytic anemia; and bleeding diathesis have been described previously as presenting features.\textsuperscript{6}

**CASE REPORT**

The proband was born at 36 weeks’ gestation via normal vaginal delivery. The pregnancy had been complicated by vaginal bleeding at 12 weeks’ gestation. The mother reported that all of her other antenatal tests were within reference limits. The infant’s anthropometric measurements were appropriate for gestational age. She was born to nonconsanguineous white parents. There was no family history of jaundice, bleeding disorders, or FTT.

There was no evidence of hyperbilirubinemia or intrahepatic cholestasis in the neonatal period. No elevations of tyrosine and galactose levels were detected by newborn screening.

At 2 months of age, failure to gain weight was noted. The infant was diagnosed clinically with gastroesophageal reflux and treated with ranitidine and metoclopramide for 3 weeks without clinical benefit. Her parents reported good oral intake of formula.

The infant presented at 7 months of age with an episode of bleeding associated with a decrease in her hemoglobin level to 2 g/dL. This incident was treated with fresh-frozen plasma and packed red blood cells.

One month later, at 8 months of age, she was admitted to a different hospital with oozing of blood from her mouth. She had a mildly elevated protime (PT) (15.4 seconds) and a low albumin level (3.2 g/dL). Her hemoglobin level dropped to 7 g/dL, which required another

![Figure 1](https://i.imgur.com/3Q5Z5Q.png)

**FIGURE 1**

Growth chart (length). NCHS indicates National Center for Health Statistics.
transfusion of packed red blood cells. She had normal liver transaminase levels at this point. A computed tomography scan of her abdomen and pelvis revealed no arteriovenous malformations. Bronchoscopy and esophagogastroduodenoscopy did not identify the source of the blood loss.

She was readmitted 1 month later, at 9 months of age, with hematochezia and profuse bleeding from her rectum. Her hemoglobin level was 5.6 g/dL, and she required additional transfusions and fresh-frozen plasma. Her PT was elevated at 19.4 seconds, and her partial thromboplastin time was elevated at 36.7 seconds.

The workup for her bleeding diathesis revealed an abnormal clotting profile with a low factor 2 (44% of control) and low factor 5 (35% of control). We failed to find a source of bleeding with repeat imaging, including angiography and a Meckel scan. She was noted on this admission to have a very low albumin level (2.8 g/dL), although her liver transaminase levels remained within the reference range. Severe growth deceleration was also noted (see Figs 1 and 2).

An extensive workup for FTT (Table 1) was essentially normal except for the plasma amino acids, which on hospital day 7 showed nonspecific elevations in several amino acids (see Table 2, plasma amino acid 1).

Given similarities with previously reported presentations, there was a concern for Munchausen syndrome by proxy7,8 and child protective services became involved.

The infant was given an observed calorie challenge using a free amino acid formula (120 kcal/kg [502 kJ/kg] per day: 41% fat, 12% protein, 47% carbohydrate) for 12 days. This intake is significantly in excess of the dietary reference intakes for a typical 9-month-old (81 kcal/kg [340.2 kJ/kg] per day: 39% fat, 6% protein, 45%
Despite this formula, she failed to gain weight.

On hospital day 20, a repeat plasma amino acid assay revealed a modest and isolated elevation of citrulline with no detected argininosuccinic acid (see Table 2, plasma amino acid 2). She had a serum lactate level within the reference range (1.4 mmol/L) and a normal urinary organic acid analysis. The result of an assay for argininosuccinate synthetase activity in cultured fibroblasts was within the reference range (4.1 nmol/min per mg protein [reference range: 0.8–3.8 nmol/min per mg protein]). The patient had ammonia (8 μmol/L [reference range: <50 μmol/L]) and urine orotic acid (0.4 mg/mL [reference range: 0.3–2.82 mg/mL]) levels within the reference ranges. A subsequent amino acid analysis was performed on hospital day 31. This test revealed a biochemical profile suggestive of NICCD (elevated citrulline, threonine, and methionine), although no elevations of tyrosine and arginine were observed (see Table 2, plasma amino acid 3).

On the basis of the experience of Imamura et al in adults with citrin deficiency, a high-protein/low-carbohydrate diet was initiated. This 6-kg infant was given 4 g/kg per day of protein, 9.3 g/kg per day of carbohydrate, and 6 g/kg per day of fat (113 kcal/kg [475 kJ/kg] per day; 15% protein; 35% carbohydrate, 50% lipids). The infant immediately began gaining weight, and energy intake was reduced to 103 kcal/kg (433 kJ/kg) per day. Full correction of growth velocity and weight was subsequently seen (see Figs 1 and 2). Plasma amino acids and the coagulation profile normalized (see Table 2, plasma amino acid 4). At 20 months of age she had a reduced protein intake (11% protein, 35% carbohydrate, and 54% fat). Her growth and weight became static. Furthermore, her mother reported some minor rectal bleeding. Her linear growth velocity resumed with no additional bleeding episodes when the protein was again increased to 15%.

**Molecular Studies**

Western blotting performed on the patient’s fibroblasts showed the absence of citrin protein with a normal adenosine triphosphate synthase band. Complementary DNA (cDNA) was prepared by using total RNA extracted from cultured skin fibroblasts, and sequencing of the coding regions was performed as described previously. This test revealed a deletion of exon 3 when sequenced in both directions.

**Discussion**

FTT is a common pediatric problem. In most cases, the cause is thought to be insufficient energy intake or constitutional small size. Conversely, inborn errors of urea cycle metabolism present classically with an acute encephalopathy or neurocognitive dysfunction or as a result of abnormal newborn screening test results. Citrin deficiency has been reported previously in Southeast Asian populations as a result of an abnormal newborn screen, liver failure, or (later on in life) with a neuropsychiatric presentation.

Our proband, the first of European ancestry reported, presented with FTT and without evidence of intrahepatic cholestasis at 6 months of age. This alternative presentation expands the clinical spectrum of citrin deficiency. Moreover, it illustrates the importance of considering rare diseases when children exhibit persistent FTT that does not respond to adequate energy intake.

It is interesting to note that our patient had a normal

---

**TABLE 1** Workup of FTT

<table>
<thead>
<tr>
<th>Test</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal sweat chloride level</td>
<td></td>
</tr>
<tr>
<td>Normal fat-soluble vitamin levels</td>
<td></td>
</tr>
<tr>
<td>Low albumin level (2.8 g/dL)</td>
<td></td>
</tr>
<tr>
<td>Normal thyroid studies</td>
<td></td>
</tr>
<tr>
<td>Normal colonoscopy and esophagogastroduodenoscopy results</td>
<td></td>
</tr>
<tr>
<td>HIV negative</td>
<td></td>
</tr>
<tr>
<td>Normal immunoglobulins</td>
<td></td>
</tr>
<tr>
<td>Erythrocyte sedimentation rate:</td>
<td>5 mm/hr</td>
</tr>
<tr>
<td>C-reactive protein:</td>
<td>&lt;0.3 mg/dL</td>
</tr>
<tr>
<td>Normal viral hepatitis panel</td>
<td></td>
</tr>
<tr>
<td>Normal chromosome analysis</td>
<td></td>
</tr>
<tr>
<td>Normal stool α-1-antitrypsin level</td>
<td></td>
</tr>
<tr>
<td>Normal fecal fats</td>
<td></td>
</tr>
<tr>
<td>Normal cholesterol and triglyceride levels</td>
<td></td>
</tr>
</tbody>
</table>

---

**TABLE 2** Results of Selected Plasma Amino Acids

<table>
<thead>
<tr>
<th>Plasma Amino Acid</th>
<th>Reference Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citrulline, μmol</td>
<td>2–41</td>
</tr>
<tr>
<td>Threonine, μmol</td>
<td>20–210</td>
</tr>
<tr>
<td>Serine, μmol</td>
<td>56–188</td>
</tr>
<tr>
<td>Methionine, μmol</td>
<td>9–45</td>
</tr>
<tr>
<td>Tyrosine, μmol</td>
<td>20–96</td>
</tr>
<tr>
<td>Arginine, μmol</td>
<td>42–132</td>
</tr>
<tr>
<td>Alanine, μmol</td>
<td>148–420</td>
</tr>
<tr>
<td>Ornithine, μmol</td>
<td>5–129</td>
</tr>
<tr>
<td>Glutamine, μmol</td>
<td>238–842</td>
</tr>
<tr>
<td>Argininosuccinate, μmol</td>
<td>0</td>
</tr>
</tbody>
</table>
Citrulline level at a time of low energy intake because she was on intravenous dextrose after a gastrointestinal bleed. It can be hypothesized that the protein load resulting from this bleed may have provided a source of arginine and/or aspartate, which could correct the metabolic defect because loss of citrin causes deficiency of aspartate in the cytosol. In addition, by restricting the carbohydrate intake, the increase in the cytosolic reduced nicotinamide-adenine dinucleotide/oxidized nicotinamide-adenine dinucleotide ratio would be reverted and a potential derangement of lipid metabolism and depletion of cytosolic aspartate would be prevented. Furthermore, the failure of growth after reduction in protein intake with continued carbohydrate restriction suggests the importance of maintaining a high-protein, low-carbohydrate diet for these patients. The observed decompensation illustrates the importance of ongoing health surveillance.

We hypothesize that the failure of liver synthetic dysfunction, as verified by the low albumin levels, is the cause of the bleeding disorder seen in citrin deficiency. Expansion of the newborn screening program can be expected to identify more cases, which would require understanding of this emerging metabolic phenotype.

ACKNOWLEDGMENTS
This study was supported in part by Grant-in-Aid for Scientific Research (B) 16390100 from the Japan Society for the Promotion of Science and by Grant for Child Health and Development 17–2 from the Ministry of Health, Labor, and Welfare of Japan.

We thank W. E. O’Brien, PhD (Biochemical Genetics Laboratory, Baylor College of Medicine) for plasma amino acid analysis and assistance with interpretation and Suzanne D’Souza, RD, LD (Texas Children’s Hospital) for assistance with dietary calculations and dietary management of this patient.

REFERENCES
12. Saheki T, Kobayashi K, Mitochondrial aspartate glutamate carrier (citrin) deficiency as the cause of adult-onset type II citrullinemia (CTLN2) and idiopathic neonatal hepatitis (NICCD). J Hum Genet. 2002;47:333–341
Sinus Bradycardia After Intravenous Pulse Methylprednisolone

Jonathan D. Akikusa, MBBSab, Brian M. Feldman, MDab,c, Gil J. Gross, MDab,d, Earl D. Silverman, MDab,e, Rayfel Schneider, MBBcha,b

Divisions of aRheumatology and cCardiology, Hospital For Sick Children, Toronto, Ontario, Canada; Departments of bPediatrics, cHealth Policy Management and Evaluation and Population Health Sciences, and eImmunology, University of Toronto, Toronto, Ontario, Canada

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

ABSTRACT

High-dose intravenous pulse methylprednisolone is an important therapeutic modality for many autoimmune conditions in both children and adults. Adverse effects of this therapy include hypertension, hyperglycemia, and, in children, behavioral changes. Cardiac rhythm disturbances, both tachyarrhythmias and bradyarrhythmias, have been reported in adults but much less commonly in children. Here we report our experience over a 6-month period with 5 children with rheumatic diseases who developed sinus bradycardia during consecutive daily therapy with intravenous pulse methylprednisolone. Reductions in resting heart rate of between 35% and 50% of baseline were observed in each case. All patients were asymptomatic, and all recovered spontaneously over a variable period of time after cessation of pulse therapy. Sinus bradycardia after repeated administration of high-dose pulse methylprednisolone in children may be more common than previously appreciated.

CASE REPORTS

CASE 1. A previously well 12-year-old boy was prescribed a 3-day course of high-dose pulse methylprednisolone therapy (1.0 g daily) for the treatment of Wegener granulomatosis. Ten hours after the second dose, ~35 hours after the first, in the early hours of the morning, the patient developed persistent resting bradycardia with a heart rate of 35 beats per minute (bpm) (previous resting heart rate: 95 bpm; Fig 1A). His blood pressure was normal, and when roused from sleep his heart rate increased transiently into the normal range. An electrocardiogram (ECG) revealed a sinus bradycardia with prolongation of the rate-corrected QT interval (QTc) (0.58 seconds [reference: <0.44 seconds]). Serum ele-
FIGURE 1
Heart rate (black line) and mean blood pressure (gray line) over time of cases 1 (A) and 2 (B), plotted as 15-point moving averages. Vertical lines indicate the timing of intravenous pulse steroid infusions, and horizontal dashed lines indicate the age-specific 5th and 95th centiles for heart rate.
trolyte and creatinine levels were within reference range. No specific therapy was given; however, the planned third dose of methylprednisolone was omitted, and high-dose oral corticosteroids were commenced instead. His resting bradycardia resolved over 5 days. A follow-up ECG was normal.

CASE 2. A previously well 6-year-old boy was prescribed a 3-day course of high-dose pulse methylprednisolone therapy (750 mg daily) for the initial treatment of juvenile dermatomyositis. Ten hours after receiving his third pulse, ~50 hours after the first dose, in the early hours of the morning, the patient developed persistent resting bradycardia with a heart rate of 45 bpm (previous resting heart rate: 93 bpm; Fig 1B). He was asymptomatic. An ECG performed at the time revealed sinus bradycardia with a normal QTc. Serum electrolyte and creatinine levels were within reference range. No specific therapy was given, and the bradycardia resolved over 8 days.

Table 1 outlines the significant features of the 2 described cases and those of the 3 other children who developed bradycardia during pulse methylprednisolone therapy at our institution. All had an adequate nutritional state. All were treated using the same protocol (30 mg/kg methylprednisolone to a maximum of 1000 mg diluted in 100 mL normal saline or 5% dextrose/0.2% normal saline infused over 30–60 minutes) and all had continuous monitoring of vital signs before, during, and after the infusions. Each patient had had >1 dose at the time of onset of the bradycardia, which occurred at a median of 50 hours after the first infusion. Recorded heart rates at the nadir of bradycardia in each patient were below the 5th percentile for age. In all the cases, bradycardia was most prominent at rest or during sleep. Although heart rates increased when the patients were roused through the day, they were usually at the lower limit of normal for age and decreased again as they fell asleep. One child had a blood pressure above the 99th centile for age at the time his bradycardia was noted. Blood pressure in the other 4 patients was normal. Electrocardiography revealed sinus bradycardia in all the cases. One patient was found to have a prolonged QTc. Four children had electrolyte levels checked at the time the bradycardia was first noted; one had a low serum calcium level, and another had borderline-low serum potassium and magnesium levels. In both cases, bradycardia persisted despite correction of the observed abnormalities. The bradycardia was asymptomatic in all patients and resolved without specific drug therapy over 3 to 10 days.

**DISCUSSION**

High-dose intravenous pulse methylprednisolone has become an important therapeutic modality for clinicians treating diseases in which rapid control of immune-mediated processes is required. Arrhythmias have been reported to occur in anywhere from 1% to 82% of patients receiving such therapy. Those most com-

| TABLE 1 | Summary of the Clinical Features and Investigations of the 5 Patients |
|---------------------|---------------------|---------------------|---------------------|---------------------|
| Parameter | Patient No. | 1 | 2 | 3 | 4 | 5 |
| Age/gender (M/F) | | 12 y/M | 6 y/M | 4 y/M | 0.7 y/M | 14 y/F |
| Indication for pulse steroid therapy | | Wegener granulomatosis | JDM | MAS (systemic JIA) | MAS (Kawasaki Disease) | Relapsing polychondritis |
| Steroid dose used | | 1.0 g | 750 mg | 600 mg | 255 mg | 1.0 g |
| No. of infusions before bradycardia onset | | 2 | 3 | 2 | 3 | 3 |
| Time to bradycardia onset from first infusion | | 35 h | 50 h | 24 h | 60 h | 50 h |
| Blood pressure at bradycardia onset | | 104/78 mm Hg | 80/52 mm Hg | 123/79 mm Hg | 88/54 mm Hg | 110/72 mm Hg |
| Baseline heart rate | | 92 bpm | 93 bpm | 80 bpm | 137 bpm | 82 bpm |
| ECG rhythm | | Sinus bradycardia | Sinus bradycardia | Sinus bradycardia | Sinus bradycardia | Sinus bradycardia |
| Abnormal serum electrolyte levels (of Na⁺, K⁺, Cl⁻, Ca²⁺, Mg²⁺, albumin) | | 0.58 s | 0.44 s | None | K⁺: 3.4 mmol/L³ | None |
| | | | | | Mg²⁺: 0.56 mmol/Lb | Albumin: 31 g/L |
| | | | | | Ca²⁺: 2.0 mmol/Lc | Not assessed |
| Reported symptoms | | None | None | None | None | None |
| Drug therapy administered | | None | None | None | None | None |
| Duration of bradycardia | | > 72 h | 8 d | 72 h | 10 d | > 72 h |
| Outcome/mean heart rate in 24 h before discharge | | Recovered/68 bpm | Recovered/75 bpm | Recovered/86 bpm | Recovered/107 bpm | Recovered/60 bpm |

JDM indicates juvenile dermatomyositis; MAS, macrophage-activation syndrome; JIA, juvenile idiopathic arthritis.

³ Reference: 3.5–5.0 mmol/L.

b Reference: 0.7–0.95 mmol/L.

c Corrected serum Ca²⁺ level (reference: 2.17–2.44 mmol/L).

e780 AKIKUSA et al
monly described are sinus bradycardia, atrial fibrillation/ flutter, and ventricular tachycardia. 10–13,22,23 The majority of reports have been in adults; however, arrhythmias have occurred in pediatric patients. 16–21 Although most frequently reported in association with intravenous methylprednisolone at a dose of 30 mg/kg (maximum of 1.0 g), arrhythmias may also occur at lower doses 20,24,25 and with administration by the intramuscular or even oral route. 19,22 They have been reported after both single and consecutive daily doses, and their onset may occur as early as during administration of the methylprednisolone to as late as several days afterward. 23,26–28 The late development of arrhythmias may make it difficult to identify methylprednisolone as the cause if the association is not appreciated. The mean time to development of bradycardia from initiation of therapy in our cases was 44 hours (range: 24–60 hours).

Symptoms resulting from arrhythmias in reported cases have ranged from none to palpitations, loss of consciousness, and cardiac arrest. 11,13,19,20,22,23 Interventions have varied from simple observation to administration of chronotropic or antiarrhythmic agents 11,28,29 to temporary cardiac pacing. 19 The reported duration of arrhythmias has varied from hours to days. All our patients were asymptomatic and continued to have bradycardia for at least 72 hours after it was first noted.

The mechanisms underlying the development of arrhythmias in association with high-dose steroids are unknown. In animal studies, high-dose methylprednisolone has significant effects on cardiovascular physiology that may be mediated both by direct action on the myocardial cell membrane and via alterations in cardiovascular sensitivity to catecholamines. 10,31 In humans, intravenous methylprednisolone alters the stimulation threshold of myocardial cells and, when given in pulse doses, alters serum potassium and the urinary excretion of both potassium and sodium; these changes might conceivably alter electrolyte shifts across myocardial cell membranes. 13,12,13 Although 2 of our patients had abnormal electrolyte levels at the time their arrhythmias were detected, correction of the abnormalities did not have any immediate effect on their bradycardia. Alternatively, it is possible that pulse methylprednisolone–induced changes in sodium and water physiology result in expansion of plasma volume, triggering a reflex bradycardia by activation of low-pressure baroreceptors. 34 Because this was a retrospective study, we were unable to explore this possibility further in our patients. Other proposed mechanisms for development of arrhythmias in association with high-dose steroid therapy include reaction to excipients in the steroid preparations or the presence of a “predisposed” myocardium. 16 One of our patients was found to have a prolonged QTc at the time of his bradycardia. Whether this represents an effect of pulse methylprednisolone (which has not previously been reported), an underlying electrophysiological defect, or a combination of both is unclear. Carriers of certain long QT syndrome gene mutations or polymorphisms may manifest a prolonged QTc only when exposed to a predisposing factor(s). 35 Hypertension is a recognized adverse effect of pulse methylprednisolone, 8 and baroreceptor-mediated reflex heart rate reduction is a potential explanation for bradycardia in this setting. All but 1 of our patients were normotensive at the time of their bradycardia, a finding similar to that of a previous report that indicated a lack of consistent change in blood pressure in adult patients with steroid-induced bradycardia. 28 Finally, it is possible that the activity of the underlying inflammatory diseases in our patients, with associated hyperdynamic circulation, contributed to the baseline heart rate seen in our patients and that reduction in inflammation as a result of treatment was responsible for the observed decrease in heart rate. Although this may be true to some extent, a sustained heart rate below the 5th percentile in this situation would not be expected. Furthermore, in each case the observed bradycardia was observed to resolve after the cessation of pulse methylprednisolone, which makes it likely to have been a “real” effect and not simply a reduction in inflammatory response. It is likely that arrhythmias after pulse methylprednisolone therapy have a multifactorial origin. Our current practice when an arrhythmia is noted during pulse methylprednisolone therapy is to obtain an ECG, check serum sodium, potassium, calcium, magnesium, and albumin levels, and correct any electrolyte abnormalities identified. Inpatients have continuous cardiac monitoring.

CONCLUSIONS
Bradycardia, particularly sinus bradycardia, may occur as an adverse effect of intravenous methylprednisolone pulse therapy administered on consecutive days in children. Reductions in heart rate are usually delayed in relation to the steroid infusion and frequently not apparent until the second or third day of therapy. Although their duration may be prolonged, typically continuing for several days after the cessation of pulse methylprednisolone therapy, symptoms are rare. Spontaneous resolution can be expected.

REFERENCES
5. Tanaka H, Tateyama T, Waga S. Methylprednisolone pulse
Scurvy and Rickets Masked by Chronic Neurologic Illness: Revisiting “Psychologic Malnutrition”

James McCallum Noble, MD, Arthur Mandel, MD, PhD, Marc C. Patterson, MD, FRACP

Departments of *Neurology and †Pediatrics, Morgan Stanley Children’s Hospital of New York and the Neurological Institute of New York, Columbia University Medical Center, New York, New York

Financial Disclosure: Dr Patterson served on the advisory boards of and/or served as a consultant for Actelion Pharmaceuticals, Inc, Amicus Therapeutics, and Stem Cells, Inc (no honoraria were retained). The other authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT
The North American epidemic of overeating, combined with a sedentary lifestyle, has led to a growing prevalence of obesity, diabetes, and the “metabolic syndrome” in children. Excessive caloric intake does not imply adequate nutrition, and vitamin-deficiency syndromes still occur in some American children. Here we describe cases of scurvy and vitamin D deficiency in 2 children with cognitive disorders. Thorough dietary histories suggested the diagnosis in each patient and, had they been obtained at presentation, would likely have obviated invasive diagnostic workup, unnecessary stress to the patients and their families, and significant functional disability. Overnutrition and malnutrition may coexist, particularly among those with abnormal cognition or autistic spectrum disorders. Classic nutritional deficiencies must not be omitted from the differential diagnosis. A comprehensive dietary history and screening for vitamin deficiencies in at-risk children are important aspects of preventive health care and are essential for prompt diagnosis and treatment.

MORE THAN 50% of adults in the United States are overweight or obese, and similar trends are emerging in children. Suggested risk factors for development of pediatric obesity include lack of physical activity, excessive television/video viewing, sugar-sweetened drink consumption, large portion sizes, consumption of low amounts of fruits and vegetables, and not breast-feeding.1 Figures from the Centers for Disease Control and Prevention in 2002 indicated increasing rates of obesity, with 16.1% of 12- to 19-year-old children being overweight (defined as those ≥95th percentile of the gender-specific BMI-for-age growth chart) and an additional 14.8% of the same population designated as at risk for becoming overweight; this trend was only slightly more common in boys than girls.2 A review of obese children and adolescents presenting to our hospital’s clinics revealed high rates of hypertension (11%–19%), hypercholesterolemia (18%), impaired glucose tolerance (7%), and metabolic syndrome (14%).3

Excess caloric intake does not imply adequate nutrition or exclude vitamin deficiencies, which, although seemingly rare, still occur in urban and suburban communities.4-6 Although Talbot7 contended in 1963 in Pediatrics that in daily practice, “psychologic malnutrition” (a term coined for poorly understood behavioral disorders) would replace easily treatable diseases such as scurvy and rickets, 2 cases of symptomatic vitamin deficiency that presented to our neurologic consultation service within 1 week suggest that certain pediatric neuropsychiatric disorders have allowed these diseases to persist, often going unrecognized for prolonged periods of time. Here we review these cases and recent studies that have suggested that, within certain populations, perhaps we should focus on dietary deficiencies as much as excess.

CASE REPORTS

CASE 1. A 5-year-old white boy from suburban New York City presented with listlessness, behavioral regression, and an antalgic gait. He was diagnosed with pervasive developmental disorder at 18 months of age. He had no physical limitations and was highly skilled in age-appropriate activities. He had an antalgic gait and was noted to have a marasmic build. His dietary history was consistent with a sedentary lifestyle, excessive video game playing, and a low-protein diet. He consumed many sugary drinks and few vegetables. He was started on vitamin D3 and Smolka’s regimen for treatment of scurvy and his condition improved. He was discharged and has continued to improve with vitamin D3 treatment. He makes slow but steady progress in all areas.

Key Words: scurvy, autistic spectrum disorder, vitamin D deficiency, malnutrition, hypocalcemia

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1071
doi:10.1542/peds.2006-1071

Accepted for publication Sep 22, 2006
Address correspondence to James McCallum Noble, MD, Department of Neurology, Neurological Institute of New York, Columbia University Medical Center, 710 W 168th St, New York, NY 10032.
E-mail: jn2054@cumc.columbia.edu
PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
priate video games. His socialization improved after en-
rolling in school at 5 years of age. His verbal expression
remained limited; he spoke in single words or brief
phrases. Six months before admission, he was enrolled
in a special school where his program included physical,
occupational, and speech therapy and applied behavioral
analysis.

Six weeks before admission, his teachers reported
diminished energy and withdrawal. One week later, he
was less playful and lost interest in climbing and run-
ning. While ascending a slight incline, the patient fell
onto his buttocks, cried inconsolably, and walked with a
right-sided limp thereafter. Over the next few days, he
complained of right leg pains and frequently asked to be
carried. His behavior regressed further, and he con-
stantly requested a pacifier. He lost interest in his favor-
ite computer games. He stopped talking and communi-
cated with inarticulate groans and by pointing at desired
objects. He subsequently stopped walking and took to
bed.

His orthopedist and neurologist found a limping an-
taligic gait; his regressive behavior was attributed to a
nonspecific musculoskeletal disorder.

Because of concern that he would regress further as
an inpatient, hospitalization was deferred. Outpatient
workup revealed normal blood chemistry and peripheral
blood counts. His erythrocyte sedimentation rate was 20
mm/hour (reference range: 0–15 mm/hour) at the ini-
tial evaluation and 44 mm/hour 1 week later. Radi-
ographs revealed a questionable right fibular fracture that
was not confirmed on a technetium 99m bone scan. He
had anorexia and lost 8 pounds over the next 3 weeks.

The patient was admitted to our center for evaluation
of worsening pain, fatigue, and failure to thrive. He was
bedbound at the time of admission.

He took no medications. His mother and 12-year-old
sister both had learning difficulties; a maternal uncle had
features of autism. The family declined genetic investi-
gation. The patient’s maternal grandfather died of brain
atrophy–proven Creutzfeldt-Jakob disease at the age of
60.

Examination showed a slender boy lying in bed, suck-
ing on his pacifier, whose hips and knees were always
flexed. His weight was 15 kg (5th percentile), height was
105 cm (25th percentile), and BMI was 13.6 kg/m²
(<5th percentile). He was drowsy but easily roused and
remained awake through the examination. There was a
nonblanching micropapular petechial rash on the exten-
sor surfaces of his left arm and right thigh, and gradually
enlarging ecchymoses were noted at sites of attempted
blood draws. He had no gingival bleeding. Each of his 3
middle fingers had horizontal subungual linear ecchy-
moses, each equidistant from the nail bed. No finger
biting had been witnessed by his family. His hair was
normal. Attempts to passively extend his knees and hips
provoked active flexion of his hips and knees and inar-
ticate cries. There were no joint effusions, crepitus, or
enlargement. Muscle palpation and range-of-motion
testing of his ankles, arms, neck, hands, and fingers were
unremarkable. He had limited spinal flexion and main-
tained an extended posture. Examination of his cranial
nerves, upper extremities, reflexes, and coordination
was unremarkable, as was the general medical exami-
nation. A diffuse musculoskeletal problem was sus-
pected, with the differential diagnosis including juvenile
rheumatoid arthritis, hemolytic malignancy, and
Goldbloom syndrome. An exhaustive diagnostic workup
was pursued. Table 1 summarizes the patient’s serial
laboratory data.

Radiographs revealed a radiolucent area in the left
distal radial metaphysis and a possible periosteal reaction
of the superior articular process of the L3 vertebral body.
MRI of the abdomen, pelvis, entire spine, and both legs
revealed abnormal bone marrow signal (hypointense on
T1-weighted and hyperintense on T2-weighted images)
in both distal femora and proximal tibiae with periosteal
reaction. There was abnormal marrow signal within
multiple bilateral transverse processes and pedicles of
the lumbar spine, suggesting an infiltrative process such
as leukemia or lymphoma. A full-body technetium 99m-
diphosphonate bone scan was unremarkable. Bone mar-
row biopsy revealed mildly hypocellular marrow with
megakaryocytic hyperplasia and atypia and relative ery-
thropoietic hyperplasia, without evidence of acute leukemia.
There was no evidence of acute leukemia by flow-cyto-
metric analysis. MRI of the complete neuraxis revealed
slightly prominent lateral ventricles and sulci for age.
Results of cerebrospinal fluid analysis were normal.

On the ninth hospital day, his vitamin C level (from
serum collected on admission) was reported as undetect-
able, which confirmed the diagnosis of scurvy. The assay
was suggested after review of the bone marrow biopsy
and observation of his progressive ecchymoses and gin-
gival bleeding, which began during hospitalization.

Retrospective review of his dietary history was reveal-
ing. The patient was breastfed for the first 6 months and
supplemented with nocturnal infant formula once forti-
fied infant foods were introduced. By 12 months of age
he ate table foods including pasta and corn, but he never
liked cereals or other common “finger foods.” He had
limited citrus fruits in his diet but ate bananas and was
supplemented with multivitamins until 2 years of age.
At this point his dietary behavior changed, and he pro-
gressively restricted his diet. By 3½ years of age, his daily
diet consisted only of dry Honeycomb cereal at breakfast,
strawberry yogurt for lunch, plus Goldfish crackers, Ritz
crackers, and Ritz cheese and cracker sandwiches as
snacks. For dinner he only ate generic grocery store
brand ice cream (vanilla plus chocolate sauce or vanilla
and chocolate combination), served in a small Dixie-
style cup. Once weekly he ate garlic and marinara pasta
at a local restaurant. By 1 year before admission he no
longer ate Honeycomb cereal but was substituting crackers for breakfast. Six months later he stopped eating the restaurant pasta. Unknown to his parents, he did not eat the yogurt packed for his lunch once he enrolled in school 6 months before admission, and he refused to eat it on weekends thereafter. Thus, for 6 months before admission his diet was restricted to crackers, ice cream, and water.

Discussion

References exist in ancient literature, but the first well-described cases of scurvy date to the era of oceanic (as contrasted to coastal) exploration by sailing vessels when sailors had a restricted diet that excluded perishables (including citrus fruits) for periods of months to years. Historical accounts of those afflicted with scurvy and the first attempts at treatment bear a brief review centuries later for their insightful descriptions and analyses. The first formal account of scurvy was by Jacques de Vitry in his “Historie des Croisades” in 1210:

“A large number of men in our army were attacked by a certain pestilence, against which the doctors could not find any remedy in their art. A sudden pain seized the feet and legs; immediately afterward the gums and teeth were attacked by a sort of gangrene, so the patient could not eat any more. Then the bones of the legs became horribly black, and so, after having suffered continued pain . . . a large number of Christians went to rest on the bosom of the Lord.”

Three centuries later (in 1536) Jacques Cartier attempted the first treatment of a similar blight with tree bark and sap. The description appeared in Richard Hakluyt’s The Principall Navigations (1600), appropriately entitled: “How by the grace of god we had notice of a

---

**TABLE 1 Laboratory Values for Case 1**

<table>
<thead>
<tr>
<th>Test</th>
<th>Hospital Day</th>
<th>Reference Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>White blood cells, 10 x 10^3/L</td>
<td>8.7</td>
<td>4.5–13.5</td>
</tr>
<tr>
<td>Red blood cells, 10 x 10^6/L</td>
<td>4.39</td>
<td>3.90–5.30</td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td>10.9</td>
<td>11.5–13.5</td>
</tr>
<tr>
<td>Hematocrit, %</td>
<td>33.1</td>
<td>34.0–40.0</td>
</tr>
<tr>
<td>Mean corpuscular volume, fl.</td>
<td>75.4</td>
<td>75.0–90.0</td>
</tr>
<tr>
<td>Mean corpuscular hemoglobin, pg</td>
<td>24.8</td>
<td>26.7–31.9</td>
</tr>
<tr>
<td>Mean corpuscular hemoglobin conc, g/dL</td>
<td>32.9</td>
<td>32.3–35.9</td>
</tr>
<tr>
<td>Red cell distribution width, %</td>
<td>12.5</td>
<td>&lt;14.5</td>
</tr>
<tr>
<td>Platelets, 10 x 10^9/L</td>
<td>439</td>
<td>165–415</td>
</tr>
<tr>
<td>Mean platelet volume, fl.</td>
<td>10.6</td>
<td>9.00–12.95</td>
</tr>
<tr>
<td>Neutrophils, %</td>
<td>86</td>
<td>35–60</td>
</tr>
<tr>
<td>Lymphocytes, %</td>
<td>14</td>
<td>3–8</td>
</tr>
<tr>
<td>Monocytes, %</td>
<td>0</td>
<td>0–4</td>
</tr>
<tr>
<td>Eosinophils, %</td>
<td>0</td>
<td>0–2</td>
</tr>
<tr>
<td>Basophils, %</td>
<td>0</td>
<td>0.8–2.0</td>
</tr>
<tr>
<td>Reticulocytes, %</td>
<td>3.5</td>
<td>0.8–2.0</td>
</tr>
<tr>
<td>Na, mM/L</td>
<td>135</td>
<td>136–146</td>
</tr>
<tr>
<td>K, mM/L</td>
<td>4.0</td>
<td>3.6–5.0</td>
</tr>
<tr>
<td>Cl, mM/L</td>
<td>102</td>
<td>102–109</td>
</tr>
<tr>
<td>CO₂, mM/L</td>
<td>23</td>
<td>24–35</td>
</tr>
<tr>
<td>Serum urea nitrogen, mg/dL</td>
<td>7</td>
<td>7–20</td>
</tr>
<tr>
<td>Glucose, mg/dL</td>
<td>103</td>
<td>70–105</td>
</tr>
<tr>
<td>Ca, mg/dL</td>
<td>9.4</td>
<td>8.7–10.0</td>
</tr>
<tr>
<td>Total protein, g/dL</td>
<td>6.3</td>
<td>6.7–8.6</td>
</tr>
<tr>
<td>Albumin, g/dL</td>
<td>3.1</td>
<td>4.1–5.3</td>
</tr>
<tr>
<td>Total bilirubin, mg/dL</td>
<td>0.7</td>
<td>0.30–1.30</td>
</tr>
<tr>
<td>Direct bilirubin, mg/dL</td>
<td>0.1</td>
<td>0.04–0.38</td>
</tr>
<tr>
<td>Aspartate aminotransferase, U/L</td>
<td>35</td>
<td>12–38</td>
</tr>
<tr>
<td>Alanine aminotransferase, U/L</td>
<td>25</td>
<td>7–41</td>
</tr>
<tr>
<td>Alkaline phosphatase, U/L</td>
<td>78</td>
<td>33–96</td>
</tr>
<tr>
<td>Erythrocyte sedimentation rate, mm/h</td>
<td>30</td>
<td>1–15</td>
</tr>
<tr>
<td>C-reactive protein, mg/L</td>
<td>5.81</td>
<td>&lt;3.0</td>
</tr>
<tr>
<td>Lactate dehydrogenase, U/L</td>
<td>187</td>
<td>115–221</td>
</tr>
<tr>
<td>Creatine kinase, U/L</td>
<td>53</td>
<td>115–221</td>
</tr>
<tr>
<td>Fe, μg/dL</td>
<td>20</td>
<td>41–141</td>
</tr>
<tr>
<td>Total iron-binding capacity, μg/dL</td>
<td>216</td>
<td>267–388</td>
</tr>
<tr>
<td>Serum folate, ng/mL</td>
<td>19</td>
<td>5.4–18.0</td>
</tr>
<tr>
<td>Vitamin B₁₂, pg/mL</td>
<td>887</td>
<td>279–996</td>
</tr>
</tbody>
</table>
Lind. His classic description begins:

the treatment of scurvy was reported in 1753 by James the maner how to use it." The first randomized trial of certaine tree, whereby we all recovered our health: and for supper, barley and raisins, rice and currants, sago and wine, or the like.

Two of these . . . had each two oranges and one lemon given them every day. . . . They continued but 6 days under this course, having consumed the quantity that could be spared.

The consequence was, that the most sudden and visible good effects were perceived from the use of oranges and lemons; one of those who had taken them, being at the end of 6 days fit for duty. . . . The other was the best recovered of any in his condition; and being now deemed pretty well, as appointed nurse to the rest of the sick. . . .

As I shall occasion elsewhere to take notice of the effects of other medicines in this disease, I shall here only observe, that the result of all my experiments was, that oranges and lemons were the most effectual remedies for this distemper at sea. . . .

It may now be proper to confirm the efficacy of these fruits by the experience of others.

Notwithstanding this discovery and continuing investigations, the decline of scurvy among British sailors did not begin until citrus consumption was mandated by the Merchant Shipping Amendment Act of 1867.

Data from the US Census in 1910 indicated the rarity of the disease even then with 68 (0.0084%) of 805 412 reported deaths; this was a more common cause of death than plague, leprosy, glanders, anthrax, rabies, and chronic occupational poisonings other than lead and than plague, leprosy, glanders, anthrax, rabies, and chronic lead poisoning.

Scurvy today is usually found among the elderly, the poor, alcoholics, and those restricted to macrobiotic diets or other diets that contain <10 mg/day of vitamin C. The US Department of Agriculture recommended daily allowance of vitamin C is 90 mg. Vitamin C is necessary as an antioxidant, for the synthesis of carnitine, collagen, and norepinephrine, and in prostaglandin metabolism. Manifestations of scurvy are largely related to underlying small hemorrhages, abnormal mitochondrial metabolism, and abnormal collagen synthesis and may manifest as bleeding gums and hemorrhoses, arthralgias, malaise, and weakness.

Autistic children typically restrict their diet and prefer foods of smooth texture (ie, pureed foods). Food-avoidant behaviors among these children can also be related to food presentation or specific utensil use. A review of the dietary patterns of 175 autistic children suggested that a taste “habituation” to sweet and salty foods was characteristic; high-glycemic-index foods were consistently preferred over others. Review of the literature reveals 23 case reports of scurvy in children with behaviorally restricted diets (see Table 2), including 2 children with autism and another with an unspecified behavioral problem, 1 with trisomy 21, 1 with static encephalopathy after pertussis-associated encephalitis, 2 with cerebral palsy, and 5 with global developmental delay; 7 otherwise reportedly normal children, including 1 morbidly obese teenager and another with a pathologic grief reaction, had scurvy in association with unusual nonfaddist diets and isolated food-avoidant behaviors. Recurrent themes in most of these reports were prolonged diagnostic workups, inadequate initial dietary history, diets similar to those described above, and authors’ impressions of the uniqueness of each presentation. Cases of childhood scurvy after infancy have been reported among the poor or those with parentally determined diets, including 6 obese children in a series of 28 children described in Thailand.

Our patient displayed a pattern of taste habituation to sweet and salty foods and avoidance of bitter and sour foods. His diet contained no vitamin C in the 6 months before presentation when he discontinued his only previous source, strawberry yogurt. Small helpings of raw strawberries (1 cup = 98 mg of vitamin C) can give a complete supply of the daily recommended value of vitamin C for an adult. Review of nutritional information among the popular yogurt brands including Breyers, Yoplait, and Dannon revealed that only Dannon strawberry yogurt contained vitamin C (8% of the daily value based on an adult 2000-calorie diet). Some fortified children’s cereals, including our patient’s favorite (Honeycomb), contain no vitamin C (citrus juices are considered another part of the typical breakfast) but served as a likely source of his other vitamin requirements. Aside from elemental minerals, the crackers described contain no significant vitamins.

Vitamin C supplementation was followed by improvement in range of motion in his legs, behavior, and pain control. A gastrostomy tube was placed for adequate caloric and vitamin intake. Two months later, he had returned to school and gained 12 lb. Six months after his hospitalization he had a 20-lb weight gain, had made a friend in school, and, although his behavioral problems persisted, his language had improved to 2- to 3-word phrases.

**Case 2.** A 17-year-old boy presented with seizures and tetany. He is a resident of Washington Heights, New York City.
York City, of black and Dominican ancestry, with a history of sickle trait and mild asthma that required occasional albuterol inhalers. He is mildly mentally retarded and attended special school with “mainstreaming” for some classes. His cognitive and social function has been stable. His mother exhibits signs of mild mental retardation and had adult-onset Graves’ disease.

The patient watched movies until 5 AM on the morning of admission. After 3 hours of sleep he awoke, and his mother was alerted by a cry from his room. She found him unresponsive with whole-body stiffening followed by limb-jerking and eye deviation. The convolution ended after 2 minutes, but he remained confused for another 15 minutes. After awakening, he complained of cramping leg pain and was brought to the hospital.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Age</th>
<th>Gender</th>
<th>Associated Diagnosis</th>
<th>Described Diet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beynon and Evans (1951)</td>
<td>23 mo</td>
<td>F</td>
<td>“Backward in development”</td>
<td>Bread, milk, corn flakes; refused orange juice and all vegetables</td>
</tr>
<tr>
<td></td>
<td>3½ y</td>
<td>F</td>
<td>Trisomy 21</td>
<td>Boiled milk, only virtually liquid food</td>
</tr>
<tr>
<td></td>
<td>3 y</td>
<td>F</td>
<td>Mental retardation, spastic limbs, Cooley’s anemia, “backwardness in development”</td>
<td>Only meat, porridge, and custard from a bottle; vomited orange juice and cod-liver oil</td>
</tr>
<tr>
<td></td>
<td>4½ y</td>
<td>F</td>
<td>“Backwardness” and seizures</td>
<td>Cereals, milk, puddings, gravy, bread and milk, fruit puree once monthly</td>
</tr>
<tr>
<td></td>
<td>2½ y</td>
<td>F</td>
<td>Pertussis-associated encephalitis at 1 y; subsequent static encephalopathy with spastic diplegia and optic atrophy</td>
<td>Breakfast: Benger’s food; lunch: gravy and potatoes; tea: Benger’s food; reportedly took doses of orange concentrate, unclear frequency</td>
</tr>
<tr>
<td></td>
<td>4½ y</td>
<td>M</td>
<td>“Primary amentia” diagnosed at age 3½ y</td>
<td>Cereals, custard, milk puddings, biscuits and cake, stewed fruit, bananas, tea, coffee, Ovaltine; all vegetables were vomited, regularly taken orange juice was stopped 1 y before</td>
</tr>
<tr>
<td>Lopresti et al (1964)</td>
<td>3 y</td>
<td>M</td>
<td>Hydrocephaus, developmental delay</td>
<td>Kool-Aid, junior mixed (cooked) vegetables with meat</td>
</tr>
<tr>
<td>Doulas et al (1973)</td>
<td>4 y</td>
<td>M</td>
<td>Behavioral problem</td>
<td>Only cow’s milk, bread coated with olive oil or sugar powder</td>
</tr>
<tr>
<td>Ellis et al (1984)</td>
<td>9 y</td>
<td>F</td>
<td>Normal</td>
<td>Tuna fish sandwiches without lettuce, iced tea without lemon</td>
</tr>
<tr>
<td>Shamash et al (1988)</td>
<td>2½ y</td>
<td>M</td>
<td>Mentally retarded, deaf-mute, able only to crawl</td>
<td>Only milk and soup</td>
</tr>
<tr>
<td>Gupta et al (1989)</td>
<td>8 y</td>
<td>F</td>
<td>Mental retardation</td>
<td>Preference for starchy foods and an aversion to fruits and vegetables; overall diet 500–600 total calories per day</td>
</tr>
<tr>
<td>Clark et al (1992)</td>
<td>7 y</td>
<td>M</td>
<td>Severely retarded and mentally handicapped</td>
<td>Primarily vanilla pudding</td>
</tr>
<tr>
<td>McKenna and Dawson (1993)</td>
<td>14 y</td>
<td>F</td>
<td>Morbidly obese</td>
<td>Unbalanced diet crisps, hamburger, chocolate bars, Coca-cola</td>
</tr>
<tr>
<td>Gómez-Carrasco et al (1994)</td>
<td>12 y</td>
<td>F</td>
<td>Pathologic grief reaction</td>
<td>Whole milk, yogurt, custard pudding, cookies, and pasta soup; ate no fruit, vegetables, legumes, meat, or fish</td>
</tr>
<tr>
<td>Shetty et al (1998)</td>
<td>6 y</td>
<td>M</td>
<td>Autism</td>
<td>Cookies, yogurt, whole milk, biscuits, and water; ate no fruit, vegetables, meat, or fish</td>
</tr>
<tr>
<td>Narchi and Thomas (2000)</td>
<td>18 mo</td>
<td>M</td>
<td>Normal</td>
<td>Normal diversity but preferred a diet consisting mainly of whole milk, yoghurt, biscuits, and cookies, disliked meat, fish, vegetables, and fruit</td>
</tr>
<tr>
<td>Chatripromp and Wananukul (2001)</td>
<td>8 y</td>
<td>M</td>
<td>Cerebral palsy, G6PD deficiency, kermicterus</td>
<td>Water, commercial chocolate puddings and cakes, occasional jarred semipureed foods (beef, spaghetti), and 2% milk</td>
</tr>
<tr>
<td>Weinstein et al (2001)</td>
<td>9 y</td>
<td>F</td>
<td>Moderate global developmental delay, mild facial dysmorphism, seizures</td>
<td>Macaroni and cheese, pizza, McDonald’s chicken nuggets, vanilla pudding pops, Doritos chips, water, and Dr Pepper soda</td>
</tr>
<tr>
<td>Monks et al (2002)</td>
<td>9 y</td>
<td>M</td>
<td>Autism, obsessive-compulsive traits</td>
<td>Rice, crackers, bread, and poultry</td>
</tr>
<tr>
<td>Bingham et al (2003)</td>
<td>16 y</td>
<td>M</td>
<td>Normal, 7 mo of unexplained diarrhea</td>
<td>Bread and water with some milk, lemonade, müesli bars, chocolate, and “Roll-ups” (a fruit-derived snack food for children)</td>
</tr>
<tr>
<td>Akikusa et al (2003)</td>
<td>9 y</td>
<td>M</td>
<td>Normal</td>
<td>Absence of vitamin C</td>
</tr>
<tr>
<td>Rosati et al (2005)</td>
<td>8 y</td>
<td>M</td>
<td>Perinatal cerebral hemorrhage, left hemiparesis</td>
<td></td>
</tr>
</tbody>
</table>
Examination in the emergency department revealed leg tenderness with diminished motion. He was awake and normally interactive. He was hyporeflexic. Chvostek and Trousseau signs were present. The general examination was normal. Laboratory investigations showed a calcium level of 4.5 mg/dL (reference range: 8.7–10.0 mg/dL) and ionized calcium level of 0.61 mM/L (reference range: 1.12–1.32 mM/L); he had no evidence of renal insufficiency. A 2-minute seizure observed in the emergency department was aborted with intravenous lorazepam. He maintained a frog-leg position while sedated. He experienced several episodes of tetany of the trunk and arms. Both the seizures and tetany were abolished by parenteral calcium repletion. Subsequent laboratory values revealed an appropriately high parathyroid hormone level (515 pg/mL [reference range: 8–51 pg/mL]), a slightly high phosphorous level (5.0 mg/dL [reference range: 2.5–4.3 mg/dL]), and elevated creatine kinase level (1516 U/L [reference range: 51–294 U/L]). A skeletal survey revealed no dystrophic calcification, but there was marked osteopenia and bilateral femoral neck fractures, which were treated with open reduction and internal fixation. Both 1,25-OH and 25-OH vitamin D levels were undetectable. Laboratory values 7 years before presentation showed a normal serum calcium level.

The patient had no recent history of paresthesias, leg cramping, cognitive decline, psychosis, or muscle pains. Dietary reporting was inconsistent. His mother described a poor diet consisting largely of fast foods, including his favorite food: cheese pizza. He drank 3 L of carbonated cola daily. He had no identified sources of excessive fluoride ingestion. He reported taking 3 children’s chewable vitamins nearly daily from 5 years of age. These vitamins contained 400 IU of vitamin D in each tablet. His mother did not corroborate his reported vitamin intake. He was active in sports including basketball and football throughout the year but remained indoors during the winter. Two days before admission he played in a local indoor basketball game for >1 hour without fatigue or cramping. He denied symptoms of malabsorption.

Subsequent investigations indicated depleted liver stores of fat-soluble vitamins. His serum α-tocopherol level was 3.8 mg/L (reference range: 5.5–18.0 mg/L), γ-tocopherol level was 0.7 mg/L (reference range: 0.0–6.0 mg/L), retinol level was 0.07 mg/L (reference range: 0.26–0.70 mg/L), and retinyl palmitate level was 0.00 mg/L (reference range: 0.00–0.10 mg/L). Results of coagulation studies were normal (prothrombin time: 15.0 seconds [reference range: 12.7–15.4 seconds]; international normalized ratio: 1.12). The elevated phosphate level was unexpected in the setting of hyperparathyroidism and, if not a laboratory error, may have been attributable to rhabdomyolysis provoked by tetany and seizures. Subsequent phosphate values declined to within reference levels 2 hours later, after receiving initial parenteral calcium therapy; he became hypophosphatemic by hospital day 5. Without evidence for malabsorption, the subsequent normalization of serum fat-soluble vitamin levels after supplementation strongly suggested primary nutritional deficiency as the etiology of his acquired hypocalcemia. The subsequent clinical course was complicated by avascular necrosis of the femoral heads, which required removal of his orthopedic hardware 10 months after his original presentation.

Discussion

Vitamin D deficiency is associated with several common syndromes. Rickets, a childhood form of vitamin D deficiency, was first described in Daniel Whistler’s 1644 thesis, De Morbo Puerile Anglorum Quem Patrio Idiomate Indignae Vocant the Rickets. It remains frequent in developing nations and is usually related to a combination of inadequate dietary intake and exposure to sunlight. Wharton and Bishop reviewed the epidemiology and pathophysiology of rickets including the inadequate mineralization of osteoid and other consequences of hypocalcemia, including seizures. Vitamin D deficiency in adults manifests as osteoporosis or osteopenia and is particularly prevalent in elderly patients with hip fracture.

A growing body of evidence reflects high rates of acquired hypocalcemia and hypovitaminosis D in children and adolescents. Associated risk factors include maternal antenatal vitamin D intake and sun exposure, black and Hispanic ethnicity, winter and spring seasons, BMI, low milk and high fruit juice/soft-drink consumption, and infrequent physical activity. That this patient presented toward the end of winter in a northern American city (New York City lies at latitude 40°N) may be significant. Among 307 otherwise well 11- to 18-year-olds in urban Boston (latitude 42°N), 24.1% were vitamin D deficient. In a study of 9- to 16-year-old children who presented at the end of winter to a pediatric emergency facility in Edmonton, Alberta, Canada (latitude 52°N), 69% of the boys and 35% of the girls had vitamin D insufficiency (considered to be <40 nmol/L); no patients manifested vitamin D deficiency when intake exceeded 0.45 μg/kg per day. In Helsinki, Finland (latitude 60°N), 61.8% of 178 otherwise well 14- to 16-year-old girls were vitamin D deficient during winter, which was significantly associated with low bone mineral density. The American population is gradually drifting southward, yet 50% of Americans still live north of latitude 39°N. However, winter vitamin D deficiency was found in one third of black patients aged 12 to 29 years living in the South.

Long-bone fracture attributed to acquired vitamin D deficiency in otherwise normal pubertal and adolescent
patients has not been described. This case demonstrates that careful dietary scrutiny and empiric vitamin supplementation may be necessary, even with reportedly sufficient dietary vitamin D.

CONCLUSIONS
These cases emphasize the role of comprehensive history-taking in the recognition of classic deficiency diseases. Idiosyncratic diets are common in the population, albeit more likely in patients with cognitive and communication disorders. An apparently normal diet should not preclude appropriate supplementation. Professional focus on the consequences of excess caloric intake should not lead physicians to neglect the potential for malnutrition, particularly among young and otherwise healthy children. A careful assessment of the diet and, when appropriate, screening for vitamin deficiencies should be an integral part of every child’s medical care.

ACKNOWLEDGMENTS
OGT 918–007, a phase I/II controlled trial of n-butyldexonojirimycin in adult and juvenile Niemann-Pick disease, type C, was given clinical support by Actelion Pharmaceuticals, Inc (Dr Patterson, principal investigator).

We thank Jay Selman, MD, for assistance with case 1 and Deborah M. Levy, MD, for superb diagnostic acumen.

REFERENCES
7. Talbot NB. Has psychologic malnutrition taken the place of rickets and scurvy in contemporary pediatric practice? Pediatr. 1963;31:909–918
10. Lind J. A Treatise of the Scurvy in Three Parts, Containing an Inquiry Into the Nature, Causes and Cure of That Disease Together With a Critical and Chronological View of What Has Been Published on the

Subject. Edinburgh, Scotland: Sands, Murray and Cochran for A. Kincaid and A. Donaldson; 1753

PEDIATRICS Volume 119, Number 3, March 2007 e789
38. Ritz crackers [packaging information]. East Hanover, NJ: Nabisco; 2005
39. Goldfish [packaging information]. Norwalk, CT: Pepperidge Farms; 2005