ARTICLES

Digital Childhood: Electronic Media and Technology Use Among Infants, Toddlers, and Preschoolers
Elizabeth A. Vandewater, Victoria J. Rideout, Ellen A. Wartella, Xuan Huang, June H. Lee, and Mi-suk Shim

Dyssomnias and Parasomnias in Early Childhood
Dominique Petit, Évelyne Touchette, Richard E. Tremblay, Michel Boivin, and Jacques Montplaisir

Impact of Medicaid Disenrollment on Health Care Use and Cost
Mary E. Rimsza, Richard J. Butler, and William G. Johnson

Health Insurance Across Vulnerable Ages: Patterns and Disparities From Adolescence to the Early 30s
Sally H. Adams, Paul W. Newacheck, M. Jane Park, Claire D. Brindis, and Charles E. Irwin, Jr

Psychological Functioning and Coping Among Mothers of Children With Autism: A Population-Based Study
Guillermo Montes and Jill S. Halterman

Need for and Use of Family Leave Among Parents of Children With Special Health Care Needs
Paul J. Chung, Craig F. Garfield, Marc N. Elliott, Colleen Carey, Carl Eriksson, and Mark A. Schuster

Sleep Duration and Overweight in Adolescents: Self-reported Sleep Hours Versus Time Diaries
Kristen L. Knutson and Diane S. Lauderdale

Duration of Poverty and Child Health in the Quebec Longitudinal Study of Child Development: Longitudinal Analysis of a Birth Cohort
Louise Séguin, Béatrice Nikiema, Lise Gauvin, Maria-Victoria Zunzunegui, and Qian Xu

Randomized Clinical Trial of Prevention of Hydrocephalus After Intraventricular Hemorrhage in Preterm Infants: Brain-Washing Versus Tapping Fluid
Andrew Whitelaw, David Evans, Michael Carter, Marianne Thoresen, Jolanta Wroblewska, Marek Mandera, Janusz Szwietlinski, Judith Simpson, Constantinos Hajivassiliou, Linda P. Hunt, and Ian Pople
Neurodevelopmental and Growth Outcomes of Extremely Low Birth Weight Infants Who Are Transferred From Neonatal Intensive Care Units to Level I or II Nurseries
Shabnam Lainwala, Rebecca Perritt, Kenneth Poole, Betty Vohr for the National Institute of Child Health and Human Development Neonatal Research Network

Childhood Cancer and Birthmarks in the Collaborative Perinatal Project
Kimberly J. Johnson, Logan G. Spector, Mark A. Klebanoff, and Julie A. Ross

Healing of Hymenal Injuries in Prepubertal and Adolescent Girls: A Descriptive Study
John McCann, Sheridan Miyamoto, Cathy Boyle, and Kristen Rogers

The Effect of Breastfeeding on Cardiorespiratory Risk Factors in Adult Life
Alicja R. Rudnicka, Christopher G. Owen, and David P. Strachan

Long-term Follow-up of 414 HIV-Infected Romanian Children and Adolescents Receiving Lopinavir/Ritonavir-Containing Highly Active Antiretroviral Therapy
Mark W. Kline, Sorin Rugina, Margareta Ilie, Rodica F. Matusa, Ana-Maria Schweitzer, Nancy R. Calles, and Heidi L. Schwarzwald
Pediatrics 2007; 119: e1116-e1120.

Brain Abnormalities in Patients With Hyperimmunoglobulin E Syndrome
Alexandra F. Freeman, Christina J. Collura-Burke, Nicholas J. Patronas, Lidia Stana Ilcus, Dirk Darnell, Joie Davis, Jennifer M. Puck, and Steven M. Holland

A Longitudinal Study of the Prevalence, Development, and Persistence of HIV/Sexually Transmitted Infection Risk Behaviors in Delinquent Youth: Implications for Health Care in the Community
Erin Gregory Romero, Linda A. Teplin, Gary M. McClelland, Karen M. Abram, Leah J. Welty, and Jason J. Washburn
Pediatrics 2007; 119: e1126-e1141.

Hospital-Based Directly Observed Therapy for HIV-Infected Children and Adolescents to Assess Adherence to Antiretroviral Medications
Daniel Glikman, Linda Walsh, Judy Valkenburg, P. Daisy Mangat, and John F. Marcinak

Upper-Limb Botulinum Toxin A Injection and Occupational Therapy in Children With Hemiplegic Cerebral Palsy Identified From a Population Register: A Single-Blind, Randomized, Controlled Trial
Remo N. Russo, Maria Crotty, Michelle D. Miller, Sonya Murchland, Peter Flett, and Eric Haan

Predicting Pediatric Distress During Radiation Therapy Procedures: The Role of Medical, Psychosocial, and Demographic Factors
James L. Klosky, Vida L. Tyc, Xin Tong, Dee Kumar Srivastava, Mindy Kronenberg, Alberto J. de Armendi, and Thomas E. Merchant
Exposure to Movie Smoking Among US Adolescents Aged 10 to 14 Years: A Population Estimate  
James D. Sargent, Susanne E. Tanski, and Jennifer Gibson  

Altering Portion Sizes and Eating Rate to Attenuate Gorging During a Fast Food Meal: Effects on Energy Intake  
Cara B. Ebbeling, Erica Garcia-Lago, Michael M. Leidig, Linda G. Seger-Shippee, Henry A. Feldman, and David S. Ludwig  

A Cluster-Randomized Trial of Benchmarking and Multimodal Quality Improvement to Improve Rates of Survival Free of Bronchopulmonary Dysplasia for Infants With Birth Weights of Less Than 1250 Grams  
Michele Walsh, Abbott Laptook, S. Nadya Kazzi, William A. Engle, Qing Yao, Maynard Rasmussen, Susie Buchter, Gregory Heldt, William Rhine, Rose Higgins, Kenneth Poole for the National Institute of Child Health and Human Development Neonatal Research Network  

How Reliable Is a Negative Blood Culture Result? Volume of Blood Submitted for Culture in Routine Practice in a Children's Hospital  
Thomas G. Connell, Mhisti Rele, Donna Cowley, Jim P. Buttery, and Nigel Curtis  

Effectiveness of Trimethoprim/Sulfamethoxazole for Children With Chronic Active Otitis Media: A Randomized, Placebo-Controlled Trial  
Erwin L. van der Veen, Maroeka M. Rovers, Frans W. J. Albers, Elisabeth A. M. Sanders, and Anne G. M. Schilde  

Aminoglycoside-Based Triple-Antibiotic Therapy Versus Monotherapy for Children With Ruptured Appendicitis  
Adam B. Goldin, Robert S. Sawin, Michelle M. Garrison, Danielle M. Zerr, and Dimitri A. Christakis  

Determinants of Outcomes After Head Cooling for Neonatal Encephalopathy  

Mechanisms, Clinical Presentations, Injuries, and Outcomes From Inflicted Versus Noninflicted Head Trauma During Infancy: Results of a Prospective, Multicentered, Comparative Study  
Kent P. Hymel, Kathi L. Makoroff, Antoinette L. Laskey, Mark R. Conaway, and James A. Blackman  

Developmental Outcome After Epilepsy Surgery in Infancy  
Tobias Loddenkemper, Katherine D. Holland, Lisa D. Stanford, Prakash Kotagal, William Bingaman, and Elaine Wyllie  

Therapeutic Drug Monitoring for Caffeine in Preterm Neonates: An Unnecessary Exercise?  
Girija Natarajan, Mirjana-Lulic Botica, Ronald Thomas, and Jacob V. Aranda  
Maternal Antibodies in Breast Milk Protect the Child From Enterovirus Infections
Karita Sadeharju, Mikael Knip, Suvi M. Virtanen, Erkki Savilahti, Sisko Tauriainen, Pentti Koskela, Hans K. Åkerblom, Heikki Hyöty, and the Finnish TRIGR Study Group

Adverse Associations of Infant and Child Sleep Problems and Parent Health: An Australian Population Study
Joanna Martin, Harriet Hiscock, Pollyanna Hardy, Belinda Davey, and Melissa Wake

Effects of Managed Care on Service Use and Access for Publicly Insured Children With Chronic Health Conditions
Amy Davidoff, Ian Hill, Brigette Courtot, and Emerald Adams

REVIEW ARTICLES

"Urticaria Multiforme": A Case Series and Review of Acute Annular Urticarial Hypersensitivity Syndromes in Children
Kara N. Shah, Paul J. Honig, and Albert C. Yan

Therapy for Head Lice Based on Life Cycle, Resistance, and Safety Considerations
Mark Lebwohl, Lily Clark, and Jacob Levitt

SPECIAL ARTICLES

Pain Reduction During Pediatric Immunizations: Evidence-Based Review and Recommendations
Neil L. Schechter, William T. Zempsky, Lindsey L. Cohen, Patrick J. McGrath, C. Meghan McMurtry, and Nancy S. Bright

Implementing Pay-for-Performance in the Neonatal Intensive Care Unit
Jochen Profit, John A. F. Zupancic, Jeffrey B. Gould, and Laura A. Petersen

COMMENTARIES

Healing of Hymenal Injuries: Implications for Child Health Care Professionals
Howard Dubowitz

Women in Pediatrics: Recommendations for the Future
Women Chairs of the Association of Medical School Pediatric Department Chairs

Circumcision in the Time of HIV: When Is There Enough Evidence to Revise the American Academy of Pediatrics' Policy on Circumcision?
Joseph D. Dickerman
General Recommendations on Immunization: Recommendations From the Advisory Committee on Immunization Practices
Pediatrics 2007; 119: 1008.

Inhalant Abuse
Janet F. Williams, Michael Storck, and the Committee on Substance Abuse, and and Committee on Native American Child Health

Maltreatment of Children With Disabilities
Roberta A. Hibbard, Larry W. Desch, and the Committee on Child Abuse and Neglect, and and Council on Children With Disabilities

Beyond Munchausen Syndrome by Proxy: Identification and Treatment of Child Abuse in a Medical Setting
John Stirling, Jr and and the Committee on Child Abuse and Neglect

AAP Publications Reaffirmed, January 2007

American Pediatric Society 2007 John Howland Award Recipient

LETTERS TO THE EDITOR

Movement of Bilirubin and Bilirubin Conjugates Across the Placenta
Antony F. McDonagh

Movement of Bilirubin and Bilirubin Conjugates Across the Placenta: In Reply
Francesco Raimondi, Letizia Capasso, Fiorella Migliaro, Antonia Romano, and Roberto Paludetto

Uninsured Latino Children: A Call to Action
Eugenia Garvin

Uninsured Latino Children: A Call to Action: In Reply
Glenn Flores, Sandra C. Tomany-Korman, and Milagros Abreu

Is the Binge-Drinking Glass Half Full or Half Empty?
James C. Turner
Is the Binge-Drinking Glass Half Full or Half Empty?: In Reply
Jacqueline W. Miller, Timothy S. Naimi, and Robert D. Brewer

Neonatal Blue-Light Phototherapy Could Increase the Risk of Dysplastic Nevus Development
Zsanett Csoma, Peter Hencz, Hajnalka Orvos, Lajos Kemeny, Attila Dobozy, Eva Dosa-Racz, Zsuzsanna Erdei, Dora Bartusek, and Judit Olah

Neonatal Blue-Light Phototherapy Could Increase the Risk of Dysplastic Nevus Development: In Reply
M. Jeffrey Maisels and Thomas B. Newman

Kernicterus, the Daubert Decision, and Evidence-Based Medicine
M. Jeffrey Maisels and Thomas B. Newman

Kernicterus, the Daubert Decision, and Evidence-Based Medicine: In Reply
John T. Sartore and Rebecca van Doren

Both Extremes of Arterial Carbon Dioxide Pressure and the Magnitude of Fluctuations in Arterial Carbon Dioxide Pressure Are Associated With Severe Intraventricular Hemorrhage in Preterm Infants
Jeffrey R. Kaiser

Both Extremes of Arterial Carbon Dioxide Pressure and the Magnitude of Fluctuations in Arterial Carbon Dioxide Pressure Are Associated With Severe Intraventricular Hemorrhage in Preterm Infants: In Reply
Namasivayam Ambalavanan

Deterrent to Healthy Lifestyles in Our Communities
Ediriweera B.R. Desapriya, Ian Pike, Anamaria Basic, and Sayed Subzwari

Deterrent to Healthy Lifestyles in Our Communities: In Reply
James Fisk

Autologous Cord Blood Transplantation in a Child With Acute Lymphoblastic Leukemia and Central Nervous System Relapse
Christian Urban, Wolfgang Schwinger, Martin Benesch, Petra Sovinz, Guenter Henze, and Hildegard Greinix

Autologous Cord Blood Transplantation in a Child With Acute Lymphoblastic Leukemia and Central Nervous System Relapse: In Reply
Ammar Hayani
Improved Outcomes of Extremely Low Birth Weight Infants
Margot Ahronovich, Ida Sue Baron, and Fern Litman

ERRATA

ERRATA

EXPERIENCE & REASON

Importance of the Clinical Recognition of Loeys-Dietz Syndrome in the Neonatal Period
Anji T. Yetman, Rebecca S. Beroukhim, Dunbar D. Ivy, and David Manchester

Diagnosis of Common Variable Immunodeficiency in a Patient With Primary Ciliary Dyskinesia

Brugada Syndrome Masquerading as Febrile Seizures

Epstein-Barr Virus–Induced Hemophagocytic Lymphohistiocytosis and X-Linked Lymphoproliferative Disease: A Mimicker of Sepsis in the Pediatric Intensive Care Unit
Matthew Mischler, Geoffrey M. Fleming, Thomas P. Shanley, Lisa Madden, John Levine, Valerie Castle, Alexandra H. Filipovich, and Timothy T. Cornell
Digital Childhood: Electronic Media and Technology Use Among Infants, Toddlers, and Preschoolers

Elizabeth A. Vandewater, PhD\textsuperscript{a,b,c,d}, Victoria J. Rideout, MA\textsuperscript{e}, Ellen A. Wartella, PhD\textsuperscript{d,f}, Xuan Huang, MA\textsuperscript{b,c,d}, June H. Lee, PhD\textsuperscript{b,c,d,g}, Mi-suk Shim, PhD\textsuperscript{b,c,d}

\textsuperscript{a}Population Research Center, \textsuperscript{b}Human Development and Family Sciences, \textsuperscript{c}Children and Technology Research Center, and \textsuperscript{d}Children’s Digital Media Center, University of Texas at Austin, Austin, Texas; \textsuperscript{e}Henry J. Kaiser Family Foundation, Menlo Park, California; \textsuperscript{f}Department of Psychology, University of California-Riverside, Riverside, California; \textsuperscript{g}Sesame Workshop, New York, New York

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

ABSTRACT

OBJECTIVES. The objectives of this study were to describe media access and use among US children aged 0 to 6, to assess how many young children fall within the American Academy of Pediatrics media-use guidelines, to identify demographic and family factors predicting American Academy of Pediatrics media-use guideline adherence, and to assess the relation of guideline adherence to reading and playing outdoors.

METHODS. Data from a representative sample of parents of children aged 0 to 6 ($N = 1051$) in 2005 were used. Descriptive analyses, logistic regression, and multivariate analyses of covariance were used as appropriate.

RESULTS. On a typical day, 75\% of children watched television and 32\% watched videos/DVDs, for approximately 1 hour and 20 minutes, on average. New media are also making inroads with young children: 27\% of 5- to 6-year-olds used a computer (for 50 minutes on average) on a typical day. Many young children (one fifth of 0- to 2-year-olds and more than one third of 3- to 6-year-olds) also have a television in their bedroom. The most common reason given was that it frees up other televisions in the house so that other family members can watch their own shows (54\%). The majority of children aged 3 to 6 fell within the American Academy of Pediatrics guidelines, but 70\% of 0- to 2-year-olds did not.

CONCLUSIONS. This study is the first to provide comprehensive information regarding the extent of media use among young children in the United States. These children are growing up in a media-saturated environment with almost universal access to television, and a striking number have a television in their bedroom. Media and technology are here to stay and are virtually guaranteed to play an ever-increasing role in daily life, even among the very young. Additional research on their developmental impact is crucial to public health.
Recent years have seen an explosion in electronic media marketed directly at the very youngest children in our society: a booming market of videotapes and DVDs aimed at infants aged 1 to 18 months, the launching of the entire television networks specifically targeting children as young as 12 months, the development of a variety of handheld video game players for preschoolers, and a multimillion-dollar industry selling computer games for children as young as 9 months. Despite this plethora of new media aimed at the very young, little is known about young children’s use of such media or the impact of such media use on children’s development. Few existing studies focus on the media use of children who are younger than 5.

The striking dearth of empirically based knowledge stands in stark contrast to popular, policy, parental, and academic interest in the impact of media on young children. The American Academy of Pediatrics (AAP) has recommended that pediatricians advise parents to avoid television-viewing entirely for children who are younger than 2 years and to limit the viewing time of older children to no more than 2 hours a day. The AAP recommendation states that pediatricians should “disourage television-viewing for children younger than 2 years and encourage more interactive activities that will promote proper brain development, such as talking, playing, singing, and reading together.” These guidelines also specify that television has no place in children’s bedrooms. At the time these recommendations were made, no comprehensive study of the actual amount of electronic media of various types that are used by very young children existed. Likewise, the extent of television placement in young children’s bedrooms was unknown at the population level. Some extremely limited data on the use of television (but not other electronic media) by young children exist in the National Longitudinal Study of Youth, but by and large, even existing data on television use have focused on children at least 2 or 3 years of age, with the vast majority focusing on children of school age or older.

This study represents an effort to address the distinct paucity of population-level data regarding the media use of a more current generation of children aged 6 months to 6 years, the very cohort that has experienced much of the marketing and explosion in availability of new media designed for their age group. The data used here were collected by the Kaiser Family Foundation in 2005 and represent the most recent and in-depth information available to date on the extent of access to and use of media among very young children in America. Our goals in this article were fourfold: (1) to describe the media exposure and media use of the current generation of very young children, (2) to assess how many of these young children fall within and outside the AAP guidelines regarding young children’s media use, (3) to identify important demographic and family predictors of whether children fall within or outside the AAP guidelines for television-viewing, and (4) to examine whether children who fall within versus outside those guidelines differ with respect to 2 activities that often are thought to be related to television-viewing: time spent reading and time spent playing outdoors.

METHODS

The survey was designed in consultation with media experts who were convened by the Kaiser Family Foundation and the University of Texas Children’s Digital Media Center. The initial questionnaire was refined through pretesting, and the data were collected through telephone interviews by Princeton Data Source from September 12 through November 21, 2005.

Sample

The survey included 1051 parents of children who were aged 6 months to 6 years. Participants were selected by random-digit telephone dialing and completed a telephone survey. Interviewers made up to 10 attempts to contact each sampled telephone number; the response rate was 33%. Calls were staggered over times of day and days of the week. For each household that was eligible, interviewers asked to speak with the parent who spent the most time with the target child. If neither parent spent more time with the child, then 1 was randomly chosen for the interview. The vast majority (81%) of respondents were mothers. The study was approved by the Institutional Review Board at the University of Texas at Austin. Six children were excluded from this study because they were missing essential age information. Therefore, the sample used comprised information from 1045 parents (for children aged 0–2, n = 412; for children aged 3–4, n = 304; for children aged 5–6, n = 329).

The sample was weighted to yield nationally representative estimates in statistical analyses. Sixty percent of the sample were identified as non-Hispanic white, 14% as black, 20% as Hispanic/Latino, and the remainder as other. With respect to income, 6% of the sample reported annual incomes of $10 000 or less, 10% of $10 000 to $19 999, 13% of $20 000 to $29 999, 21% of $30 000 to $49 000, 18% of $50 000 to $74 999, 11% of $75 000 to $99 999, and 11% of $100 000 or more (10% were either unsure or declined to answer). Forty-one percent of mother’s worked full time, 17% worked part time, and 38% were not in the paid-labor force. Seventy-six percent of children were in 2-parent family homes, and 23% were in single-parent family homes; the sample mapped extremely well to demographic characteristics of the US population on the basis of Current Population Survey estimates for 2005 (http://pubdb3.census.gov/macro/032006/hhinc/toc.htm).
Media Measures

Household Media Ownership and Access
Respondents were asked to indicate whether they had a television in the household, whether they had cable or satellite television, and the number of televisions in the household. They were also asked to indicate in terms of presence (do you have?) and number (how many?) for the following types of media: portable DVDs, VCR or DVD players, video game consoles (eg, X-box, PlayStation), handheld video game players (eg, GameBoy), and computers (desktops and laptops). They were also asked whether they had Internet access on their home computer and whether the access was high speed. Finally, respondents were asked whether their child had a television in the bedroom or not and, if they did, whether that television gets some cable or satellite channels, gets only regular channels, is used only for watching videos or playing games, or is not currently working or used. Respondents were also asked about the presence of the following in their child’s bedroom: VCR/DVD, video game console, computer, and Internet access.

Media and Technology Use
Respondents were asked to report whether their child did the following on the previous day: watched television, watched a video or DVD, played video games on a console, played hand-held video games, played computer games, used the computer for something other than games, read electronic books, and listened to music. Respondents were then asked to report on the amount of time their children spent using these various media on the previous day. The response categories were 5 minutes, 15 minutes, 30 minutes, 45 minutes, 1 hour, and 1.5 hours and up, in half-hour increments. For these questions, when the previous day was atypical, parents were asked to think about the last day they followed their typical routine.

AAP Guideline Groups
To assess whether children fell within the AAP guidelines (no television for children younger than 2 years, no more than 2 hours per day after that), we combined reports on the amount of time that children spent watching television and videos or DVDs the previous day. For children who were aged birth to 2, those who did not watch television or videos/DVDs were classified as lying within the guidelines, whereas those who watched any television were classified as lying outside the guidelines (0: within guideline; 1: outside guideline). Children who were older than 2 years were classified as within the guidelines when they watched <2 hours of television or videos/DVDs the previous day and outside the guidelines when they watched ≥2 hours (0: within guideline; 1: outside guideline).

Factors That Predicted Membership in AAP Guideline Groups
We identified 2 overarching categories of variables as possible important predictors of whether children would fall within the AAP Committee on Public Education guidelines: (1) demographic factors and (2) family media access and orientation factors.

Demographic Factors
Respondents reported on demographic variables such as their ethnicity, their educational level and employment status, and their spouse’s educational level and employment status. Demographic factors included parent’s ethnicity, indicated by 4 dummy variables: black (0: no; 1: yes), Latino (0: no; 1: yes), white (0: no; 1: yes), and other (0: no; 1: yes); maternal education, measured using a 7-point scale that ranged from 1 (none or grades 1–8) to 7 (postgraduate training or professional schooling after college); maternal employment, indicated by 3 dummy variables: full-time employment (0: no; 1: yes), part-time employment (0: no; 1: yes), and nonemployment (0: no; 1: yes); family income, measured on a scale ranging from 1 (less than $10 000) to 7 ($100 000 or more); family structure: (0: 2-parent family; 1: 1-parent family); and child gender (1: girl; 0: boy).

Family Media Access and Orientation Factors
Media access and orientation factors included whether there was a television in the child’s bedroom (0: no; 1: yes) or a video game player in the child’s bedroom (0: no; 1: yes); whether the child was in a constant television household (0: no; 1: yes) in which the parent reported that the television was on always or most of the time, even when no one was watching; parents’ report of their perception that television mostly helps children’s learning (0: does not help; 1: helps); parents’ report of their perception that television mostly hurts children’s learning (0: does not hurt; 1: hurts); parents’ report of whether they had rules about the amount of time they allowed their children to spend watching television (0: no; 1: yes); parents’ report of whether they had rules about the types of programs they allowed their children to watch on television (0: no; 1: yes).

Time Spent Reading and Playing Outdoors
Reading and playing outdoors are 2 activities that generally are deemed to be developmentally crucial for healthy development in childhood. Because time spent with electronic media is commonly believed to interfere with these 2 important activities, we examined mean differences in reading and playing outdoors between children who fell within versus outside the AAP media guidelines. As with media use, parents were asked to report the amount of time their child spent reading or being read to and playing outside. Response categories
were 5 minutes, 15 minutes, 30 minutes, 45 minutes, 1 hour, and 1.5 hours and up, in half-hour increments.

Analysis Plan
Our first 2 goals were to describe the media exposure and media use of the current generation of very young children and assess how many of these young children fell within and outside the AAP guidelines regarding young children’s media use. These issues were examined via descriptive statistics.

Our next 2 goals were to identify important predictors of whether children fell within or outside the AAP guidelines for television-viewing and to examine developmental outcomes that were associated with falling within or outside those guidelines. To ascertain factors that were related to whether children fell within the AAP guidelines, we conducted logistic regressions using 2 broad categories of predictors to predict AAP guideline group membership: demographic factors (parents’ ethnicity, maternal education, maternal employment, family income, family structure, and child gender) and family media access and orientation factors (television in child’s bedroom, video games in child’s bedroom, constant television, parental perceptions of television, and parental rules about television).

To analyze the relations between being in the AAP-guideline groups and time spent reading or being read to and playing outside, we conducted multivariate analyses of covariance (MANCOVA) to protect against type I error. In these analyses, all of the demographic and media access and orientation factors were included as covariates.

Partly because children who are aged 6 months to 6 years reflect very different developmental stages and partly because the AAP recommendation focuses on children who are younger than 2 years, all analyses were conducted separately for 3 age groups: (1) 0 to 2 years; (2) 3 to 4 years; and (3) 5 to 6 years.

In the logistic regression and MANCOVA models, 4% of the data points were missing. Multiple imputation methods were used to address missing data. The basic idea behind multiple imputation methods is to create several complete data sets, each of which has missing values suitably imputed. Each of these imputed data sets is then analyzed independently and estimates of parameters of interest are averaged across data sets, resulting in a single robust estimate. Standard errors are computed according to Rubin’s rules, devised to allow for the between- and within-imputation components of variation in the parameter estimates.

Stata’s user-written program ICE (Imputation by Chained Equations) was adopted to implement the multiple imputation. Ten data sets were imputed for the logistic regression and MANCOVA models, respectively. Because the Stata-based ICE program does not support analysis of variance, the imputed data sets were converted to SAS (SAS, Chicago, IL) and SAS was then used to conduct MANCOVAs on the imputed data sets.

RESULTS
Household Media Ownership and Access
Table 1 presents descriptive statistics regarding household media ownership and access. This table clearly shows that very young children are growing up in a media-saturated environment. Television ownership among these families was nearly universal (99.5% for children aged 0–2, 98.4% for children aged 3–4, 98% for children aged 5–6). Among households with a television set present, the average number of working televisions was 2.53, 2.78, and 2.98 for the 3 age groups, respectively. VCR/DVD players were also nearly universal, computer ownership was ~80%, and approximately half of the households had a video game console. Given their age, a fairly large proportion of these very young children had televisions in their bedrooms. This was true of almost one fifth (18%) of the 0- to 2-year-olds and more than one third of 3- to 4-year-olds (43%) and 5- to 6-year-olds (37%). Half (51%) of all children with a television in their bedroom had access to cable or satellite channels. In a similar vein, of the children with a television in their bedroom, extremely small percentages of parents (0%, 0.3%, and 0.9%, respectively) reported that this television was not currently working or being used. Taken together, this indicates that when young children have a television in their bedroom, it is generally a working television, actively in use.

A series of follow-up questions were asked of parents to ascertain their reasons for putting a television in their young child’s bedroom. These responses are presented in Table 2. The most common reason that parents named for having a television in their child’s bedroom was that it frees up other televisions in the house so that other family members can watch their own shows, followed by it keeps the child occupied so that the parent can do things around the house, it helps the child fall asleep, and it is used as a reward for good behavior.

Media and Technology Use
Table 3 presents descriptive statistics regarding media and technology use among very young children. Most of the children watched television on a “typical day” (~63% of 0- to 2-year-olds, 82% of 3- to 4-year-olds, and 78% of 5- to 6-year-olds). On that typical day, children of all ages spent an average of 1 hour and 19 minutes watching television. Approximately one third of children watched videos or DVDs on a typical day, and they spent an average of 1 hour and 18 minutes viewing them. Very few children (2%, 13%, and 16% for ages

PEDIATRICS Volume 119, Number 5, May 2007 e1009
0–2, 3–4, and 5–6, respectively) played video games (console or hand-held) on a typical day. However, when they did play, they spent an average of 55 minutes playing. Approximately 4% of 0- to 2-year-olds, 20% of 3- to 4-year-olds, and 27% of 5- to 6-year-olds used the computer on a typical day (those who did spent an average of 50 minutes at the keyboard).

Categorizing Children According to AAP Media Guidelines

The sample that was used to examine the predictors and outcomes associated with falling within versus outside the AAP media-use guidelines was restricted to those with complete data on AAP guideline variables (n = 1044). For the other variables in the model, multiple imputation was conducted to impute the missing values. Therefore, for these analyses, there were 412 children aged 0 to 2, 303 children aged 3 to 4, and 329 children aged 5 to 6.

Among 0- to 2-year-olds, only 32% (n = 131) fell within the AAP media guidelines for children of this age (no television), whereas 68% (n = 281) fell outside the guidelines. Among the older groups of children, 56% (n = 170) of 3- to 4-year-olds watched for 2 hours or less as recommended by the AAP, and 70% (n = 230) of 5- to 6-year-olds watched no more than this recommended limit.

### TABLE 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>0- to 2-y-Olds (n = 412)</th>
<th>3- to 4-y-Olds (n = 304)</th>
<th>5- to 6-y-Olds (n = 329)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Television in household, %</td>
<td>99.5</td>
<td>98.4</td>
<td>98</td>
</tr>
<tr>
<td>No. of televisions in household</td>
<td>2.53 (1.12)</td>
<td>2.78 (1.14)</td>
<td>2.98 (1.29)</td>
</tr>
<tr>
<td>Cable or satellite television in household, %</td>
<td>79</td>
<td>79</td>
<td>84</td>
</tr>
<tr>
<td>TiVo/other digital video recorder on television, %</td>
<td>21</td>
<td>18</td>
<td>23</td>
</tr>
<tr>
<td>Surround sound on television, %</td>
<td>41</td>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td>Portable DVD in household, %</td>
<td>30</td>
<td>36</td>
<td>32</td>
</tr>
<tr>
<td>No. of portable DVDs</td>
<td>1.49 (0.85)</td>
<td>1.27 (0.59)</td>
<td>1.44 (0.73)</td>
</tr>
<tr>
<td>VCR or DVD in household, %</td>
<td>94</td>
<td>93</td>
<td>93</td>
</tr>
<tr>
<td>No. of VCRs or DVDs</td>
<td>2.08 (1.09)</td>
<td>2.31 (1.24)</td>
<td>2.24 (1.29)</td>
</tr>
<tr>
<td>Video game console in household, %</td>
<td>45</td>
<td>49</td>
<td>52</td>
</tr>
<tr>
<td>No. of video game consoles</td>
<td>1.41 (0.92)</td>
<td>1.36 (0.69)</td>
<td>1.53 (0.97)</td>
</tr>
<tr>
<td>Hand-held video game in household, %</td>
<td>21</td>
<td>29</td>
<td>33</td>
</tr>
<tr>
<td>No. of hand held video games</td>
<td>1.53 (0.78)</td>
<td>1.75 (1.01)</td>
<td>1.66 (0.92)</td>
</tr>
<tr>
<td>Computer in household, %</td>
<td>80</td>
<td>79</td>
<td>82</td>
</tr>
<tr>
<td>No. of computers</td>
<td>1.47 (0.88)</td>
<td>1.45 (0.78)</td>
<td>1.72 (0.99)</td>
</tr>
<tr>
<td>Internet access on home computer, %</td>
<td>73</td>
<td>71</td>
<td>73</td>
</tr>
<tr>
<td>High-speed internet access on home computer, %</td>
<td>45</td>
<td>42</td>
<td>50</td>
</tr>
<tr>
<td>Television in child's bedroom, %</td>
<td>18</td>
<td>43</td>
<td>37</td>
</tr>
<tr>
<td>Bedroom television has some cable or satellite channels, %</td>
<td>9</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>Bedroom television has only regular channels, %</td>
<td>5</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Bedroom television only used for watching videos or play games, %</td>
<td>5</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Bedroom television not currently working or being used, %</td>
<td>0</td>
<td>0.3</td>
<td>0.9</td>
</tr>
<tr>
<td>VCR/DVD in child’s bedroom, %</td>
<td>14</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>Video game console in child’s bedroom, %</td>
<td>2</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>Computer in child’s bedroom, %</td>
<td>3</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Internet access on bedroom computer, %</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

* Among those with respective media in the household. Weighted means and SDs are reported.

### TABLE 2

<table>
<thead>
<tr>
<th>Reason</th>
<th>0- to 2-y-Olds (n = 76), %</th>
<th>3- to 4-y-Olds (n = 118), %</th>
<th>5- to 6-y-Olds (n = 121), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>So the parent and other family members can watch their own shows</td>
<td>47</td>
<td>52</td>
<td>60</td>
</tr>
<tr>
<td>It keeps the child occupied in his or her room so that the parent can do things around the house</td>
<td>42</td>
<td>42</td>
<td>31</td>
</tr>
<tr>
<td>It helps the child fall asleep</td>
<td>28</td>
<td>34</td>
<td>26</td>
</tr>
<tr>
<td>As a reward for good behavior</td>
<td>16</td>
<td>25</td>
<td>31</td>
</tr>
<tr>
<td>Stops fights between siblings</td>
<td>16</td>
<td>20</td>
<td>31</td>
</tr>
<tr>
<td>Had an extra television set that they did not want to throw out</td>
<td>24</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td>To get child to go to bed in his or her own room</td>
<td>18</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>All of child’s friends have televisions in their bedroom</td>
<td>9</td>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>

* Among parents of children who have a television in their bedroom.
TABLE 3  Descriptive Statistics: Media and Technology Use

<table>
<thead>
<tr>
<th>Parameter</th>
<th>0- to 2-y-Olds</th>
<th>3- to 4-y-Olds</th>
<th>5- to 6-y-Olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watched televisiona</td>
<td>63</td>
<td>82</td>
<td>78</td>
</tr>
<tr>
<td>Minutes of televisionb</td>
<td>75.18 (67.30)</td>
<td>89.67 (57.97)</td>
<td>73.99 (53.45)</td>
</tr>
<tr>
<td>Watched video/DVD</td>
<td>30</td>
<td>42</td>
<td>29</td>
</tr>
<tr>
<td>Minutes of video/DVD</td>
<td>67.44 (47.70)</td>
<td>86.82 (49.47)</td>
<td>78.88 (43.08)</td>
</tr>
<tr>
<td>Played video game (console/hand-held video)</td>
<td>2</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Minutes of video game (console/hand-held video)</td>
<td>51.21 (42.98)</td>
<td>60.50 (55.94)</td>
<td>51.59 (42.33)</td>
</tr>
<tr>
<td>Used computera</td>
<td>4</td>
<td>20</td>
<td>27</td>
</tr>
<tr>
<td>Minutes of computer useb</td>
<td>60.76 (49.08)</td>
<td>45.81 (33.37)</td>
<td>50.50 (40.94)</td>
</tr>
<tr>
<td>Read electronic booka</td>
<td>12</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>Minutes of reading electronic bookb</td>
<td>46.27 (49.68)</td>
<td>36.47 (22.60)</td>
<td>47.20 (48.70)</td>
</tr>
<tr>
<td>Listening to musicb</td>
<td>87</td>
<td>85</td>
<td>77</td>
</tr>
<tr>
<td>Minutes of musicb</td>
<td>63.87 (60.65)</td>
<td>57.41 (54.65)</td>
<td>51.39 (64.05)</td>
</tr>
</tbody>
</table>

a Whether children did the activity yesterday/on the last day that was a typical day. Percentages are reported.
b Weighted means and SDs are reported; among those who used the respective media.

Demographic and Family Media Factors That Predicted Whether Children Fell Outside the AAP Recommended Guidelines

Descriptive statistics for all predictors and outcomes that were associated with AAP media guideline groups are presented in Table 4. Results for logistic regressions that predicted whether children fell outside the AAP media guidelines are shown in Table 5. The logistic regression results indicate that among the youngest children (0–2), family media factors accounted for membership in AAP-guideline groups, whereas demographic factors were unrelated. Among the 2 older age groups (ages 3–4 and ages 5–6), family media factors continued to be important predictors of membership in AAP-guideline

TABLE 4  Descriptive Statistics: Predictors and Outcomes Associated With AAP Guideline Groups

<table>
<thead>
<tr>
<th>Parameter</th>
<th>0- to 2-y-Olds</th>
<th>3- to 4-y-Olds</th>
<th>5- to 6-y-Olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic factor predictors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent’s ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>13</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Latino</td>
<td>18</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>White</td>
<td>64</td>
<td>69</td>
<td>71</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Maternal employment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full time</td>
<td>39</td>
<td>37</td>
<td>47</td>
</tr>
<tr>
<td>Part time</td>
<td>18</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Not employed</td>
<td>40</td>
<td>43</td>
<td>33</td>
</tr>
<tr>
<td>Maternal educationa</td>
<td>4.34 (1.71)</td>
<td>4.51 (1.63)</td>
<td>4.35 (1.58)</td>
</tr>
<tr>
<td>Family incomeb</td>
<td>4.17 (1.75)</td>
<td>4.33 (1.71)</td>
<td>4.42 (1.64)</td>
</tr>
<tr>
<td>Family structurec</td>
<td>0.24 (0.43)</td>
<td>0.26 (0.44)</td>
<td>0.21 (0.41)</td>
</tr>
<tr>
<td>Child genderd</td>
<td>0.48 (0.50)</td>
<td>0.48 (0.50)</td>
<td>0.48 (0.50)</td>
</tr>
<tr>
<td>Media access and orientation factor predictors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Television in child’s bedroom</td>
<td>0.22 (0.41)</td>
<td>0.42 (0.50)</td>
<td>0.39 (0.49)</td>
</tr>
<tr>
<td>Video game in child’s bedroom</td>
<td>0.04 (0.20)</td>
<td>0.14 (0.34)</td>
<td>0.28 (0.45)</td>
</tr>
<tr>
<td>Constant television</td>
<td>0.34 (0.48)</td>
<td>0.34 (0.47)</td>
<td>0.30 (0.46)</td>
</tr>
<tr>
<td>Parental perception</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Television helps</td>
<td>0.39 (0.49)</td>
<td>0.48 (0.50)</td>
<td>0.41 (0.49)</td>
</tr>
<tr>
<td>Television hurts</td>
<td>0.37 (0.48)</td>
<td>0.30 (0.46)</td>
<td>0.33 (0.47)</td>
</tr>
<tr>
<td>Television time rules</td>
<td>0.52 (0.50)</td>
<td>0.60 (0.49)</td>
<td>0.71 (0.46)</td>
</tr>
<tr>
<td>Television program rules</td>
<td>0.76 (0.43)</td>
<td>0.90 (0.29)</td>
<td>0.93 (0.26)</td>
</tr>
<tr>
<td>Outcome variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minutes spent playing outside</td>
<td>65.54 (73.49)</td>
<td>95.99 (85.25)</td>
<td>94.55 (94.27)</td>
</tr>
<tr>
<td>Minutes reading/being read to</td>
<td>38.95 (43.18)</td>
<td>39.90 (34.20)</td>
<td>43.26 (44.29)</td>
</tr>
</tbody>
</table>

Data are weighted mean (SD) or percentage.
a Seven-point scale ranging from 1 (none or grades 1–8) to 7 (postgraduate training or professional schooling after college).
b Seven-point scale ranging from 1 (less than $10,000) to 7 ($100,000 or more).
c 0: 2-parent family; 1: 1-parent family.
d 0: boy; 1: girl.
e 0: no; 1: yes.
groups, with the addition of gender for 3- to 4-year-olds and being in a single-parent household for 5- to 6-year-olds.

**0- to 2-Year-Olds**

As mentioned, only family media factors predicted whether children fell outside the guidelines. Specifically, very young children with a television in their bedroom were 4 times more likely to fall outside the guidelines compared with those without a bedroom television. When parents perceived that television mostly helps learning, children were 2 times more likely to be outside the AAP guidelines (see Table 5).

**3- to 4-Year-Olds**

Both demographic and family media factors predicted whether 3- to 4-year-olds fell outside the AAP guidelines. Girls had greater odds of being outside the guidelines than boys. Three- to 4-year-old children who lived in a constant television household were 2 times more likely to fall outside the guidelines compared with those who did not (see Table 5).

**5- to 6-Year-Olds**

Among 5- to 6-year-olds, children who lived in single-parent families were more likely to be outside the guidelines than children who lived in 2-parent families. Having parents who perceived that television is mostly helpful was associated with greater odds of falling outside the guidelines, whereas having parental rules regarding television content was associated with greater likelihood of falling within the AAP guidelines (see Table 5).

**Time Spent in Other Activities**

There were no differences in time spent either with print media (reading or being read to) or playing outdoors between children who fell within versus outside the AAP media guidelines for any age group.

### TABLE 5 Logistic Regression Predicting the Likelihood of Falling Outside the AAP Recommended Television-Viewing Guidelines

<table>
<thead>
<tr>
<th>Variables</th>
<th>0- to 2-y-Olds (n = 412)</th>
<th>3- to 4-y-Olds (n = 303)</th>
<th>5- to 6-y-Olds (n = 329)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>Demographic variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent’s ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>1.09 (0.47–2.55)</td>
<td>1.02 (0.38–2.75)</td>
<td>1.02 (0.39–2.65)</td>
</tr>
<tr>
<td>Latino</td>
<td>1.05 (0.54–2.04)</td>
<td>1.71 (0.80–3.65)</td>
<td>1.29 (0.59–2.84)</td>
</tr>
<tr>
<td>Other</td>
<td>1.48 (0.45–4.89)</td>
<td>0.93 (0.23–3.83)</td>
<td>2.24 (0.57–8.80)</td>
</tr>
<tr>
<td>White</td>
<td>1.00 (—)</td>
<td>1.00 (—)</td>
<td>1.00 (—)</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>0.87 (0.35–2.18)</td>
<td>1.32 (0.49–3.57)</td>
<td>0.93 (0.31–2.80)</td>
</tr>
<tr>
<td>High school</td>
<td>1.10 (0.54–2.27)</td>
<td>0.70 (0.31–1.60)</td>
<td>1.05 (0.52–2.11)</td>
</tr>
<tr>
<td>Some college</td>
<td>0.94 (0.51–1.76)</td>
<td>0.69 (0.33–1.45)</td>
<td>1.71 (0.86–3.40)</td>
</tr>
<tr>
<td>College degree</td>
<td>0.91 (0.47–1.76)</td>
<td>0.94 (0.44–1.98)</td>
<td>0.64 (0.30–1.36)</td>
</tr>
<tr>
<td>Graduate school</td>
<td>1.00 (—)</td>
<td>1.00 (—)</td>
<td>1.00 (—)</td>
</tr>
<tr>
<td>Maternal employment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full time</td>
<td>0.82 (0.48–1.38)</td>
<td>0.69 (0.37–1.29)</td>
<td>0.79 (0.43–1.44)</td>
</tr>
<tr>
<td>Part time</td>
<td>0.73 (0.37–1.44)</td>
<td>0.75 (0.35–1.59)</td>
<td>0.56 (0.24–1.32)</td>
</tr>
<tr>
<td>Not employed</td>
<td>1.00 (—)</td>
<td>1.00 (—)</td>
<td>1.00 (—)</td>
</tr>
<tr>
<td>Family income</td>
<td>0.99 (0.83–1.19)</td>
<td>0.84 (0.68–1.04)</td>
<td>1.03 (0.83–1.27)</td>
</tr>
<tr>
<td>Family structure*</td>
<td>0.83 (0.41–1.68)</td>
<td>0.52 (0.22–1.27)</td>
<td>2.27* (1.07–4.83)</td>
</tr>
<tr>
<td>Child gender**</td>
<td>0.79 (0.49–1.27)</td>
<td>1.89* (1.09–3.27)</td>
<td>1.61 (0.94–2.78)</td>
</tr>
<tr>
<td>Family media factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Television in child’s bedroomc</td>
<td>4.45* (1.99–9.99)</td>
<td>1.45 (0.80–2.64)</td>
<td>0.71 (0.36–1.38)</td>
</tr>
<tr>
<td>Video game in child’s bedroomc</td>
<td>1.01 (0.34–2.97)</td>
<td>0.74 (0.30–1.78)</td>
<td>1.35 (0.59–3.11)</td>
</tr>
<tr>
<td>Constant televisionc</td>
<td>1.21 (0.69–2.11)</td>
<td>2.12* (1.13–3.97)</td>
<td>1.33 (0.71–2.49)</td>
</tr>
<tr>
<td>Parental perception</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Television helpsc</td>
<td>2.22* (1.23–4.00)</td>
<td>1.05 (0.52–2.14)</td>
<td>1.96* (1.05–3.68)</td>
</tr>
<tr>
<td>Television hurtsc</td>
<td>0.90 (0.49–1.66)</td>
<td>0.56 (0.24–1.27)</td>
<td>0.81 (0.38–1.72)</td>
</tr>
<tr>
<td>Television time rulesc</td>
<td>0.63 (0.31–1.26)</td>
<td>0.75 (0.41–1.38)</td>
<td>0.80 (0.42–1.52)</td>
</tr>
<tr>
<td>Television program rulesc</td>
<td>1.90 (0.94–3.85)</td>
<td>1.93 (0.65–5.71)</td>
<td>0.33* (0.12–0.90)</td>
</tr>
</tbody>
</table>

OR indicates odds ratio; CI, confidence interval.

* 0, 2-parent family; 1, 1-parent family.
** 0, boy; 1, girl.
*0, no; 1, yes.
* P < .05
* P < .01
DISCUSSION

The impetus for this study came from the striking lack of empirically based knowledge about the extent of exposure to and use of media and technology by very young children. It is interesting that at the time the AAP Committee on Public Education made its media-use recommendations,2 almost nothing was known about the extent to which very young children actually used media and technology or the extent to which young children had television in their bedrooms. This study is among the first to examine the full extent of very young children’s media use at the population level. As such, it provides the most recent information available regarding the extent of media use and media access of infants, toddlers, and preschoolers in America.

Overall, the results indicate that the AAP Committee on Public Education was certainly correct in assuming that infants, toddlers, and preschoolers are using media and using it daily. In addition, the results of this survey make it clear that very young children today are growing up in a media-saturated environment. For this cohort of children and presumably for future cohorts as well, access to and use of media have become part of the fabric of their daily lives.

Much has been made of the vast array of media and technology that are used by children today, and it is clear that they have far more choices available to them than previous generations. However, it is also true that this study indicates that children’s use of electronic media is still, by and large, dominated by television. However, when young children do play video games or use the computer, they use these media for just under an hour, which is not a trivial amount of time in a young child’s day.

To state that television and VCR ownership was virtually ubiquitous in the homes of young children does not adequately capture the extent of media saturation in these homes. More relevant is that most of them are living in homes with at least 2 of each product; nearly one quarter live in homes with ≥4 televisions. Eight in 10 children this age live in homes with access to either cable- or satellite-based channels. Taken together, this means that not only do children have many opportunities to watch televised programs, but also they have many televised programs from which to choose. Approximately half of the households had a video game console, and between one fifth (for the 0- to 2-year-olds) and one third (33% of the 5- to 6-year-olds) had access to a hand-held video game. Even more households (78%) reported having a computer, and nearly 7 in 10 of all households (69%) have Internet access.

Although none of this is perhaps surprising, we did find it surprising that many of these very young children have televisions in their own bedrooms. This was true of almost one fifth of the 0- to 2-year-olds and more than one third of the 3- to 4-year-olds and the 5- to 6-year-olds. The most common reason that parents named for having a television in their child’s bedroom was that it frees up other televisions in the house so that other family members can watch their own shows (cited by 54% of parents whose children have television in the bedroom). Other commonly cited reasons were that it keeps the child occupied so that the parent can do things around the house (38%), it helps the child fall asleep (30%), and it is used as a reward for good behavior (25%).

Whatever the reason, this trend is particularly noteworthy, not only because the AAP advises against this practice but also because there is increasing evidence that bedroom television in particular is linked to a number of poor outcomes, including academic, social, and physical activity.14–17 The question of why parents believe that putting a television in their young child’s bedroom is appropriate deserves attention by both pediatricians and scholars. The 2 most commonly cited reasons—(1) that it frees up the television so that other family members can watch their own shows and (2) that it keep children occupied—are particularly noteworthy. These responses indicate that television-viewing may be an increasingly isolated experience, even for very young children. Given the evidence that especially for young children, parental involvement and participation in viewing increases the positive impact of educational shows,7 both the prevalence and the developmental impact of viewing alone are important areas of future inquiry.

The notion that television can be used to soothe infants to sleep is also of particular interest, given that television most certainly serves as an activator of the reticular activating system, which stimulates attention and alertness. As such, parents who use the television to put their children to sleep may actually be putting them at risk for a future of sleep disorders.18,19 Although speculative at this point, the possibility of this connection (and other implications of bedroom television) is urgently in need of empirical examination.

The results also indicate that very young children are comfortable using these media on their own, particularly television. More than half (54%) of the 0- to 2-year-olds and almost all (82% and 88%, respectively) of the 3- to 4- and 5- to 6-year-olds could turn on the television by themselves, and many could also put in a video or DVD by themselves. Although many fewer of these young children could use a mouse, it is noteworthy that roughly 71% of the 5- to 6-year-olds could. Overall, it seems clear that these children will be very different from previous generations of children with respect to their comfort with technology and the extent to which they use all forms of technology in their daily lives.

Most of the very young children (0–2 years old) fell outside the guidelines recommended by the AAP (in a typical day, 68% of them watched television). This may
be at least partly because parents simply do not know about the recommendation. Eighty-five percent of parents surveyed indicated they had not talked to a pediatrician about their children’s media use at all. An earlier survey found that only 6% of parents of children who were age 2 years or younger were aware of the AAP viewing guidelines. However, approximately half (56%) of 3- to 4-year-olds and 70% of 5- to 6-year-olds fell within the AAP viewing guidelines. The reason that more of the older children fell within the guidelines may well be because for the very young children, no viewing is recommended; therefore, any viewing sets them outside the guidelines.

Different factors were implicated in whether children fell outside the AAP guidelines. Among the youngest children (ages 0–2), demographic variables were not linked to adherence to the guidelines; rather, family media factors were significant predictors. Among older children, both sets of predictors—demographic (family structure and child gender) and family media factors—were important predictors of guideline group membership. These findings suggest that the youngest children’s viewing is more directly related to their immediate environment, whereas that of older children is associated with both immediate and distal environments. Demographic predictors were particularly salient among the oldest age group (ages 5–6). Specifically, children who lived in a single-parent family were more than twice as likely to fall outside the AAP media guidelines compared with children who lived in a 2-parent family structure. Among 3- to 4-year-olds, girls were more likely than boys to fall outside the AAP guidelines.

It is noteworthy that ethnicity, family income, and parental education were not relevant to whether children followed the AAP guidelines. It could be that the widespread perception that these factors are important predictors of children’s viewing has been because the vast majority of existing research has focused on older children. Our findings in this area are in fact consistent with some other research on very young children that documented similar patterns. Taken together with findings from other studies, our findings suggest that sociodemographic factors may not start to affect children’s television-viewing until later in childhood. However, it may be that socioeconomic status and race/ethnicity are related to differences in the amount of time that young children spend using media but not specifically to whether they fall within or outside the AAP limits.

Media access and orientation factors operated differently among children of different age groups. An environment that directly exposes children to television—that is, being in a constant television household and having television in the child’s bedroom—is related to higher chances of falling outside the guidelines. Having rules about how much time children can spend with television decreased the likelihood of older children’s (ages 3–4 and 5–6) falling outside the guidelines. This likely represents the direct effect of restrictions in viewing time for these age groups. We did not find any relation between time spent viewing and time spent reading or in outdoor play among these children. These findings echo those of Vandewater et al, who similarly found no relation between television-viewing and either of these activities. It may be that, contrary to popular belief (and although children may not be reading or playing outdoors as much as we would like), media use is simply unrelated to these activities.

As indicated, the sample was recruited via random-digit-dialing techniques. Although ensuring a random sample, it should be noted that the response rate was 33%. It is widely known that response rates to telephone surveys of all kinds have noticeably lowered in the past decade or so. This study is no exception. However, confidence in the generalizability of the results is fostered in 2 ways. First, the characteristics of this sample are close to demographic characteristics from the 2005 Current Population Survey. Second, recent research has indicated that the effect of nonresponse on data is less critical than previously thought. It also should be noted that only half (50%) of parents interviewed said that they spent all or most of the day with their child. Therefore, it is possible that parents’ knowledge of their children’s media use is somewhat limited.

**CONCLUSIONS**

It is clear that in many American families, television is an almost constant presence in daily life. Among young children, this means that they will watch more television. These data highlight that even very young children are exposed to all sorts of screen and electronic media on a daily basis. This study is among the first to demonstrate empirically something that parents have perhaps known for some time: that media use is a normative part of their young child’s daily life.

Although the AAP made its media-related recommendations on the basis of concerns about the impact of media use on the very young, the question of media’s effects on children this age is just beginning to be examined. A recent study using population data that were collected in 1997 found that time spent watching television without parents was related to significant reductions in time spent interacting with parents. This, of course, was 1 of the concerns cited by the AAP Committee on Public Education when it made its recommendations.

At this point, there are more “unknowns” than “knowns” in terms of the impact of exposure to screen and electronic media on very young children’s development. There is some intriguing evidence that background television interferes with toddlers’ ability to focus on play. There is also correlational evidence of a
connection between children’s viewing and subsequent attentional problems. The question of the impact of this screen medium on very young children’s neurologic and attendant cognitive development is in urgent need of additional examination. Questions regarding the developmental consequences of bedroom media for infants and toddlers and of its impact on sleep are also in need of additional attention. Although there is clearly much to be discovered, the findings of this study provide a compelling case for the importance of additional research on the impact of electronic media on very young children’s developmental outcomes.

ACKNOWLEDGMENTS
This research was designed and analyzed in partnership with the Kaiser Family Foundation, where the project was directed by Victoria J. Rideout, vice president of the foundation and director of the Program for the Study of Entertainment Media and Health. Funding for the research was provided by the Kaiser Family Foundation, and field work was conducted by Princeton Survey Research Associates. This research was also conducted under the auspices of the Children’s Digital Media Center (funded by National Science Foundation grant BCS-0126127) directed by Drs Vandewater and Wartella, the Population Research Center at the University of Texas at Austin (funded by National Institute of Child Health and Human Development grant 5 R24 HD042849), the Children and Technology Research Center at the University of Texas at Austin, and the National Institute of Child Health and Human Development (grant 1 R01-HD40851-01 to Dr Vandewater).

REFERENCES
Dyssomnias and Parasomnias in Early Childhood

Dominique Petit, PhD*, Évelyne Touchette, MPSc*, Richard E. Tremblay, PhD*, Michel Boivin, PhDc,d, Jacques Montplaisir, MD, PhD**

*Sleep Disorders Center, Sacré-Cœur Hospital, Departments of Psychology and Psychiatry, and Research Unit on Children’s Psychosocial Maladjustment, University of Montreal, Montreal, Quebec, Canada; and Department of Psychology, Laval University, Quebec City, Quebec, Canada

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

ABSTRACT

OBJECTIVES. Our aim for this study was to determine the prevalence of dyssomnias and various parasomnias in early childhood and to describe their temporal evolution, gender differences, and correlates.

METHODS. This research is part of a longitudinal study of child development. A randomized, 3-level, stratified survey design was used to study a representative sample of infants who were born in 1997–1998 in the province of Quebec (Canada). When the children were 2.5 years of age, 1997 families agreed to be interviewed. The presence of dyssomnias or parasomnias was obtained from a self-administered questionnaire that was completed by the mother at each round of measures.

RESULTS. The percentage of children with frequent night wakings decreased steadily from 36.3% at age 2.5 to 13.2% at age 6. Similarly, the percentage of children who had difficulty falling asleep at night decreased significantly from 16.0% at ages 3.5 and 4 to 10% at age 5 and to 7.4% at age 6. The overall prevalence of each parasomnia for the period studied was as follows: somnambulism, 14.5%; sleep terrors, 39.8%; somniloquy, 84.4%; enuresis, 25.0%; bruxism, 45.6%; and rhythmic movements, 9.2%. Persistent somnambulism at age 6 was significantly correlated with sleep terrors and somniloquy. Persistent sleep terrors at age 6 were also correlated with somniloquy. Finally, persistent sleep terrors at age 6 were correlated with frequent night wakings. Separation anxiety was associated with persistent night wakings and with somnambulism, bruxism, sleep terrors, and somniloquy.

CONCLUSIONS. There is a high prevalence of night wakings and sleep-onset difficulties in preschool children. Parasomnias are highly prevalent in early childhood and are associated with separation anxiety. However, they have little impact on sleep duration.
DYSOMNIAS REFER TO difficulties with sleep that, in children, consist mainly of 2 problems: frequent night wakings (≥1 awakening per night) and difficulty in falling asleep (≥20–30 minutes).1 Large-scale epidemiologic surveys have reported that between one quarter and one third of children between the ages of 6 months and 5 years have difficulties in going to bed, falling asleep, or sleeping through the night.1–4 In toddlers, waking up during the night is strongly associated with parental behaviors surrounding the sleep period.7 Do associations with other areas of the child’s life emerge as the child gets older?

Children often experience 1 or a combination of parasomnias. The American Academy of Sleep Medicine defines parasomnias as “undesirable physical events or experiences that occur during entry into sleep, within sleep or during arousals from sleep.”8 Somnambulism and sleep terrors are classified as disorders of arousal. Somnambulism is defined as “a series of complex behaviors that are usually initiated during arousals from slow-wave sleep and culminate in walking around with an altered state of consciousness,” whereas sleep terrors are “arousals from slow-wave sleep accompanied by a cry or piercing scream and autonomic nervous system and behavioral manifestations of intense fear.” Somniloquy is defined as talking during sleep “with varying degrees of comprehensibility.” It can arise from both slow-wave sleep and rapid-eye-movement sleep. Sleep-related rhythmic-movement disorder “is characterized by repetitive, stereotyped, and rhythmic motor behaviors that occur predominantly at sleep onset and involve large muscle groups”; it includes body-rocking, head-rolling, and head-banging. Sleep enuresis “is characterized by recurrent involuntary voiding that occurs during sleep. In sleep enuresis, recurrent involuntary voiding occurs at least twice a week during sleep after 5 years of age.” Sleep-related bruxism “is an oral activity characterized by grinding or clenching teeth during sleep, usually associated with sleep arousals.”

Parasomnias are considered benign phenomena, especially in children, and do not usually have a serious impact on sleep quality or quantity. However, in more severe cases, they can result in injuries and sleep disruption or can be very disturbing for either the child or the family. Most studies on childhood parasomnias have been done in a retrospective manner (eg, adults were asked whether they had experienced parasomnias as children and, if so, at what age). One study9 that was conducted in our laboratory established the prevalence and evolution of parasomnias in children from ages 10 to 13 with retrospective questions asked to the mother for the period between ages 3 and 10. Prospective studies on parasomnias are lacking in early childhood, an important period for understanding the emergence and the development of these events.

The aims of the present study were to (1) determine the prevalence of dyssomnias and of various parasomnias (somnambulism, somniloquy, sleep terrors, enuresis, bruxism, and rhythmic movements) from 2.5 to 6 years of age, (2) describe their temporal evolution (disappearance versus new cases), (3) assess gender differences, (4) examine associations between various parasomnias and dyssomnias, and (5) study associations between dyssomnias and parasomnias and other areas of the child’s life.

METHODS

This research is part of a large longitudinal study entitled Quebec Longitudinal Study of Child Development conducted by the Quebec Institute of Statistics (Canada).10 Infants were recruited from the Quebec Master Birth Registry of the Ministry of Health and Social Services. A randomized, 3-level, stratified survey design was used to study a representative sample of infants who were born in 1997–1998 in the province of Quebec. The 3 levels were (1) geographic regions of Quebec, (2) each region subdivided in areas that were representative of the number of births in the region, and (3) number of children selected per area proportional to the number of births and to the gender ratio of this area. Families who lived in the northern part of the province of Quebec, Inuit territories, and First Nations reserves were excluded for technical reasons. Also, children with important medical problems were excluded from this cohort.

Sample Description

Of the 2940 children who were first selected, 265 were excluded, leaving 2675 who were invited to participate. Of these, 452 families refused to participate. Therefore, a total of 2223 children who were born in 1997–1998 and aged ∼5 months were included in this longitudinal study. At the second round, 2045 children who were aged ∼17 months were studied. Finally—and this is the purpose of this article—1997 families agreed to be reinterviewed when the children were 29 months (nearly 2.5 years of age), 1950 children were studied again at the age of 41 months (∼3.5 years of age), 1944 children were studied again at the age of ∼4 years, 1759 children were studied again at ∼5 years of age, and 1492 children were studied again at ∼6 years of age. Table 1 presents the number of children who were studied in each year of

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Description of the Longitudinal Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>Boys, n (%)</td>
</tr>
<tr>
<td>2.5</td>
<td>1006 (50.4)</td>
</tr>
<tr>
<td>3.5</td>
<td>979 (50.2)</td>
</tr>
<tr>
<td>4</td>
<td>977 (50.3)</td>
</tr>
<tr>
<td>5</td>
<td>875 (49.7)</td>
</tr>
<tr>
<td>6</td>
<td>734 (49.2)</td>
</tr>
</tbody>
</table>

NA indicates not applicable.
the study, the number and the proportion of boys and girls, and the number of dropouts from the previous year.

In the initial cohort, most (84.5%) children had a Canadian nonimmigrant mother; 15.5% were first-generation immigrants. The majority of the sample was white (88.4%), whereas black, Native Amerindian, Arab, and Asian children represent 3.4%, 0.3%, 2.0%, and 1.6% of the sample, respectively. Most (76.3%) mothers spoke French as a first language, 8.7% spoke English, and 15.0% had another first language.

Outcome Measures
Data were collected by a questionnaire and an interview that was conducted in English or French. First, the Self-Administered Questionnaire for the Mother, which took ~20 minutes to complete, provided information on the child's sleep duration at 6 years of age and parental behaviors around their child’s sleep period when the child was 2.5 years of age. Night wakings were defined as waking up at least once every night. Difficulty falling asleep was defined as taking ≥30 minutes to fall asleep. Dyssomnias and parasomnias (somnambulism, sleep terrors, sleeptalking, bruxism, and rhythmic movements) were assessed at the ages of 2.5, 3.5, 4, 5, and 6 years. Enuresis was evaluated at the ages of 5 and 6 years only because the minimum age criterion for a clinical diagnosis of enuresis is 5 years of age. For each parasomnia, the sample was separated into a category that comprised children who “never” had the parasomnia and a category that comprised children who had the parasomnia for each age at least occasionally.

Second, the Interviewer Completed Computerized Questionnaire, a face-to-face structured interview with the mother, provided information on her immigrant status and depression state (scores in the upper 10th percentile) when the child was 5 months of age. It also included questions on the child’s gender of the, birth weight, and socioeconomic status (SES) at 6 years of age and whether a divorce had occurred during the period of up to 6 years. Finally, this questionnaire provided information on the child’s social development at 6 years of age, more specifically on separation anxiety, hyperactivity-inattention, and aggressive behaviors. Separation anxiety was assessed from 5 questions: In the past 3 months, how often would you say your child (1) clung to adults or was too dependent, (2) did not want to sleep alone, (3) got very upset when separated from his or her parents, (4) was preoccupied by losing his or her parents, and (5) felt physical discomfort?

Hyperactivity-inattention was assessed by 10 questions: In the past 3 months, how often would you say your child (1) could not sit still, was restless, or hyperactive, (2) could not stop fidgeting, (3) was impulsive or acted without thinking, (4) had difficulty waiting for his or her turn in games, (5) could not settle down to do anything for more than a few moments, (6) could not wait when you promised something, (7) was easily distracted, had trouble sticking to any activity, (8) was unable to concentrate, (9) could not pay attention for long, and (10) was inattentive?

The presence of aggressive behaviors was evaluated by 10 questions: In the past 3 months, how often would you say your child (1) got into fights, (2) reacted in an aggressive manner when teased, (3) tried to dominate other children, (4) reacted in a aggressive manner when contradicted, (5) reacted with anger and fighting when somebody accidentally hurt him or her (eg, bumping into him or her), (6) physically attacked people, (7) hit, bit, or kicked other children, (8) reacted in an aggressive manner when something was taken away from him or her, (9) scared other children to get what he or she wanted, and (10) encouraged other children to pick on a particular child? Separation anxiety, hyperactivity-inattention, and aggressive behaviors were evaluated by examination of the percentages of children who had scores in the upper 10th percentile.

The ethics review committee of the Montreal Sacré-Cœur Hospital approved the study. It was centrally managed at the Quebec Institute of Statistics, which was responsible for data collection. Before participating in the study, all families had received detailed information by mail on the aims and procedures of the research program and had signed a consent form.

Statistical Analyses
Statistical analyses were conducted by using SPSS 10 (SPSS Inc, Chicago, IL). A complete set of data was obtained for night wakings (n = 988), sleep-onset difficulties (n = 1045), sleepwalking (n = 1085), sleep terrors (n = 1043), somniloquy (n = 1041), enuresis (n = 1137), bruxism (n = 1082), and rhythmic movements (n = 1058). Cochran Q tests assessed variations in the prevalence of each dyssomnia and parasomnia across time (age groups), then posthoc comparisons with McNemar tests pinpointed the ages at which changes had occurred. Correlations between various parasomnias and dyssomnias were assessed using Spearman correlation coefficients. χ² tests compared persistent sufferers (of a dyssomnia or parasomnia) and nonsufferers on different categorical variables (eg, low birth weight, mother’s depression, immigrant status, parental behaviors, family characteristics, behavioral measures), and t tests were used for between-group comparisons on nocturnal sleep time. The “persistent” quality was defined as still having a problem (with a dyssomnia or parasomnia) at the latest time of measure, which was at 6 years of age. For all tests conducted, the significance level was set at P < .01 to take into account the number of comparisons for each dyssomnia and parasomnia.
RESULTS

Prevalence and Evolution of Dyssomnias

Table 2 presents the prevalence of dyssomnias and parasomnias for the cohort at ages 2.5, 3.5, 4, 5, and 6 years as well as the overall prevalence for the entire study period. The overall prevalence represents the percentage of children in the sample who experienced the specific dyssomnia or parasomnia at least once during the study period. Cochran Q tests revealed age effects in the prevalence of night wakings and sleep-onset difficulties. The percentage of children with frequent night wakings (≥1 awakening every night) decreased steadily from 36.3% at 2.5 years to 13.2% at age 6; the prevalence at each age was significantly lower than that at the previous age, except between ages 3.5 and 2.5. Similarly, the percentage of children who had difficulty falling asleep at night (≥30 minutes) increased significantly from 12.2% at age 2.5 to 16.0% at ages 3.5 and 4 then decreased significantly to 10% at age 5 and to 7.4% at age 6.

For children who had frequent night wakings at some point in the study (n = 599, or 60.6% of the whole cohort), the problem was already apparent at 2.5 years for 59.9%; new cases then decreased each year to reach 2.5% at age 6 (Table 3). The offset of night wakings, however, was evenly distributed across ages, but in one fifth of those affected, the problem persisted until 6 years of age. A similar portrait emerges when sleep-onset difficulties are considered (n = 345, or 33% of the whole cohort). In most children with sleep-onset difficulties, the problem appeared at 2.5 to 3.5 years of age (69.3%) with few new cases later on. The offset of the problem was distributed across ages (with a peak at age 5) with more than one fifth of persistent cases (Table 3). Of the 108 children who had either of the dyssomnias (persistent night wakings or persistent sleep-onset difficulties), only 10 (9.3%) had from both problems.

Prevalence and Evolution of Parasomnias

The overall prevalence of each parasomnia from 2.5 to 6 years was as follows (Table 2): somnambulism, 14.5%; sleep terrors, 39.8%; somniloquy, 84.4%; enuresis, 25.0%; bruxism, 45.6%; and rhythmic movements, 9.2%. Again, Cochran Q tests yielded significant age effects in the prevalence of all categories, except somnambulism (Table 2). Posthoc McNemar tests revealed that sleep terrors were more frequent at ages 2.5, 3.5, and 4 than at ages 5 and 6. Somniloquy was more frequent at ages 2.5, 3.5, 4, 5, and 6 than at age 2.5 and at ages 4, 5, and 6 than at age 3.5. The prevalence of bruxism increased steadily with age, whereas rhythmic movements (body-rocking and head-banging) were more prevalent at age 2.5 than at all later ages. Enuresis was more frequent at age 5 than at age 6.

Two temporal progression profiles can be seen in Table 3 for parasomnias. The total n of each column is
the number of children who were afflicted with that particular parasomnia at any point in the study. Sleep terrors, somniloquy, and rhythmic movements appeared early with few new cases thereafter. In contrast, new cases of somnambulism and bruxism appeared almost equally at all ages.

**Gender Differences**

Gender differences for persistent dyssomnias and parasomnias at age 6 were also assessed. There were no differences for frequent night wakings or sleep-onset difficulties. For persistent parasomnias, significant gender differences were found for somnambulism prevalence (61.7% of boys and 38.3% of girls versus 47.0% of boys and 53.0% of girls for those without parasomnias; \( P < .01 \)) and enuresis prevalence (63.4% of boys and 36.6% of girls versus 44.7% of boys and 55.3% of girls for those without parasomnias; \( P < .001 \)), yielding a boy-to-girl ratio of ~2:1. There were no gender differences for the prevalence of sleep terrors, somniloquy, bruxism, or rhythmic movements.

**Associations Between Persistent Parasomnias and Dyssomnias**

Persistent somnambulism was correlated with persistent sleep terrors (\( r = 0.21, P < .01 \)) and persistent somniloquy (\( r = 0.16, P < .001 \)). Persistent sleep terrors were also correlated with persistent somniloquy (\( r = 0.19, P < .001 \)). Among persistent somnambulism cases, 98.5% also talked in their sleep and 41.1% had sleep terrors. Among sleep terror sufferers, 92.5% also talked in their sleep. Finally, sleep terrors were correlated with frequent night wakings (\( r = 0.31, P < .001 \)). Approximately 56.3% of sleep terror sufferers were considered as having frequent night wakings. There was no relationship between frequent night wakings and sleep-onset difficulties.

**Factors That Were Associated With Persistent Dyssomnias**

Children with persistent frequent night wakings at age 6 (~22% of those who had the problem at any point) differed from children without dyssomnias in having higher separation anxiety scores and having been put to bed already asleep in a greater proportion when they were 2.5 years of age (Table 4). Persistent sleep-onset difficulties were associated with insufficient SES; twice as many children with sleep-onset difficulty had a low family SES than did those without dyssomnias (Table 4). Surprisingly, persistent sleep-onset difficulties were not clearly associated with co-sleeping or separation anxiety; only trends were observed for those 2 factors. Both dyssomnias were associated with a lower nocturnal sleep time. Finally, neither dyssomnias was associated with hyperactivity-inattention. Only a trend for more aggressive behaviors was seen in children with sleep-onset difficulties (Table 4).

**Factors That Were Associated With Persistent Parasomnias**

Factors that were assessed for their relationship to the emergence of parasomnias appear in Table 5. Persistent somnambulism, sleep terrors, somniloquy, and bruxism all were found to be associated with separation anxiety. Persistent somnambulism was also related to a high hyperactivity-inattention score. Sleep terrors were also related to a recent divorce of the parents. As for persistent sleep bruxism, parental presence at bedtime was found

### TABLE 3  Ages of Onset and Offset for Children With a Dyssomnia or a Parasomnia

<table>
<thead>
<tr>
<th>Age, y</th>
<th>Dyssomnias</th>
<th>Parasomnias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Night Wakings</td>
<td>Sleep-Onset Difficulties</td>
</tr>
<tr>
<td></td>
<td>( n = 599 )</td>
<td>( n = 345 )</td>
</tr>
<tr>
<td>Onset</td>
<td>( n % )</td>
<td>( n % )</td>
</tr>
<tr>
<td>2.5</td>
<td>359 (59.9)</td>
<td>127 (36.8)</td>
</tr>
<tr>
<td>3.5</td>
<td>121 (20.2)</td>
<td>112 (32.5)</td>
</tr>
<tr>
<td>4.0</td>
<td>75 (12.5)</td>
<td>62 (18.0)</td>
</tr>
<tr>
<td>5.0</td>
<td>29 (4.9)</td>
<td>27 (7.8)</td>
</tr>
<tr>
<td>6.0</td>
<td>15 (2.5)</td>
<td>17 (4.9)</td>
</tr>
<tr>
<td>Offset</td>
<td>( n % )</td>
<td>( n % )</td>
</tr>
<tr>
<td>3.5</td>
<td>115 (19.2)</td>
<td>47 (13.6)</td>
</tr>
<tr>
<td>4.0</td>
<td>102 (17.0)</td>
<td>61 (17.7)</td>
</tr>
<tr>
<td>5.0</td>
<td>137 (22.9)</td>
<td>98 (28.4)</td>
</tr>
<tr>
<td>6.0</td>
<td>115 (19.2)</td>
<td>62 (18.0)</td>
</tr>
<tr>
<td>Persistent</td>
<td>130 (21.7)</td>
<td>77 (22.3)</td>
</tr>
<tr>
<td>Throughout</td>
<td>20 (3.3)</td>
<td>7 (2.0)</td>
</tr>
</tbody>
</table>

NA indicates not applicable. For each sleep problem, the \( n \) value at the top of each column indicates the number of children who experienced the problem at least once in the study period; it corresponds to the number in the “overall” row of Table 2. Age of onset was defined as the age when a given dyssomnia or parasomnia was first reported, and the age of offset was defined as the last time the dyssomnias or parasomnias were reported. Note that a dyssomnia or parasomnia may alternatively disappear and reappear. The term “persistent” signifies the number (and percentage) of children who still had the problem at 6 years of age (also includes children for whom the problem appeared at 6 years). The term “throughout” represents the number (and percentage) of children affected with the sleep problem throughout the study period.

* Parasomnias with a relatively early onset.
**TABLE 4 Factors That Were Assessed in Relation to Persistent Dyssomnias at 6 Years of Age**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Persistent Night Wakings</th>
<th>Persistent Sleep-Onset Difficulties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes (n = 130)</td>
<td>No (n = 389)</td>
</tr>
<tr>
<td>Low birth weight &lt; 2.5 kg, %</td>
<td>4.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Mother characteristic at 5 mo, %</td>
<td>11.8</td>
<td></td>
</tr>
<tr>
<td>Immigrant status</td>
<td>7.0</td>
<td>6.4</td>
</tr>
<tr>
<td>Depressive behaviors (&gt;90th percentile)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental behaviors at 29 mo, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bedtime (parental presence)</td>
<td>27.3</td>
<td>14.9</td>
</tr>
<tr>
<td>Awakenings (reactive co-sleeping)</td>
<td>17.4</td>
<td>13.0</td>
</tr>
<tr>
<td>Family characteristic at 6 y, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insufficient SES</td>
<td>11.9</td>
<td>12.3</td>
</tr>
<tr>
<td>Type, modified (divorce)</td>
<td>33.9</td>
<td>26.1</td>
</tr>
<tr>
<td>Social development at 6 y, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separation anxiety (&gt;90th percentile)</td>
<td>26.6</td>
<td>14.4</td>
</tr>
<tr>
<td>Hyperactivity-inattention (&gt;90th percentile)</td>
<td>14.8</td>
<td>11.7</td>
</tr>
<tr>
<td>Aggressive behaviors (&gt;90th percentile), %</td>
<td>12.5</td>
<td>13.3</td>
</tr>
<tr>
<td>Sleep parameters at 6 y, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nocturnal sleep time, h, min</td>
<td>10, 22 (50)</td>
<td>10, 39 (45)</td>
</tr>
</tbody>
</table>

The term “persistent” signifies that the sleep problem was still present at 6 years of age (and includes children for whom the problem appeared at 6 years). Children with a persistent sleep problem were compared with children who never had that problem (n values in the “no” categories are equal to the number of children with a complete data set for that problem minus the number of children in the overall category for that problem; see Table 2).

a Statistically significant at the threshold selected.

in a greater proportion in those with bruxism than in those without parasomnias. Persistent rhythmic movements were related to insufficient SES and depression in the mother. Enuresis could not be linked to any particular factor, except gender. Finally, none of the parasomnias covaried with low birth weight or immigrant status of the mother.

**DISCUSSION**

**Dyssomnias: Prevalence, Gender Differences, and Associations**

This study confirms in a prospective manner the high prevalence of night wakings and sleep-onset difficulties in preschool children. The incidence of both decreased steadily each year to reach 13.2% and 7.4%, respectively, at 6 years of age. It must be stressed that it is not the same children who present with the 2 types of problems; no association was found between them. Although some smaller studies have reported comorbidity of night wakings and sleep-onset difficulties in the same children,11,12,13,14 others (1 based on 5813 school-aged children)15,16 are consistent with our results in finding little or no association. As reported by other studies on children who were older than 1 year,11,12,13,14,15 no gender difference was found for either type of dyssomnia.

**Frequent Night Wakings**

In this study, frequent night wakings were associated with slightly less nocturnal sleep time. Night wakers (“signalers”) were in fact found to stay awake longer after they wake up than “self-soothers.”2 This study also shows that persistent night wakings at 6 years of age are associated with separation anxiety. However, separation anxiety in the child often stems from insecure attachment or separation issues in the mother. Scher and Blumberg16 found that 1-year-old signalers were more commonly found among mothers with high separation anxiety scores. In another study, all mothers of young children with disturbed sleep were classified as insecurely attached compared with 57% of mothers of children with normal sleep.17 In fact, it was shown with video recordings2 that all children wake up at least a few times a night. The difference between “good sleepers” and “night wakers” is often determined by whether the child decides/needs to signal the awakening. In sum, insecure attachment between mother and child will affect both (1) the child’s capacity to fall asleep alone and soothe himself or herself back to sleep if an awakening occurs without signaling it to the parents and (2) the behaviors that are adopted by the mother at the child’s bedtime and in response to his or her awakening at night.

As a matter of fact, the present study also found an association of frequent night wakings with parental behaviors of staying present until the child falls asleep (at bedtime). The same parental behaviors have been previously identified as problematic in a study of the same cohort at ages 5, 17, and 29 months; they were among the factors that were most strongly associated with the behavior of not sleeping through the night.18 As early as 1979, Anders19 stressed the importance of adopting parental behaviors that foster sleep autonomy in the very young child and showed that sleep problems in infancy covary with parenting styles at bedtime. Similarly, certain parental behaviors in response to night wakings,
such as bringing the child into the parents’ bed, will have the same effect. In support of this notion, behavioral methods of treating children’s sleep disturbances that also redress parental behaviors vis-à-vis the child’s sleep periods markedly improve or completely resolve the disturbance in 90% of cases.20

Sleep-Onset Difficulties
Persistent sleep-onset difficulties at age 6 showed an association with low family SES. This could be explained by a poor quality of sleep environment (eg, noise, sleeping arrangements), although this study did not assess this possibility. The association could also result from parental difficulties in setting limits on their children’s behavior around bedtime because of confined space or restricted resources available to them. This dyssomnia was also associated with less sleeping time. Children with sleep-onset difficulties do not seem to compensate for sleep that is lost at the beginning of the night because wake-up times often depend on social obligations (eg, school, child care because of parents’ leaving for work) and consequently are not adjustable.

Parasomnias: Prevalence, Gender Differences, and Correlates
Parasomnias are highly prevalent in preschoolers: 88% of the children in this cohort had at least 1 parasomnia during the study period. The most frequent parasomnias were somniloquy, bruxism, and sleep terrors, each affecting ≥40% of the study sample.
Somniloquy
Although considered the most frequent parasomnia, somniloquy is usually without consequences and is rarely a reason for consultation. The overall prevalence of somniloquy for preschoolers in this study (84%) is higher than what was reported in older children and adolescents. For example, an overall prevalence of 56% has been reported for children aged 3 to 13 years using mainly retrospective data.9 The association that was observed in this study between persistent somniloquy and separation anxiety is not easily explained because somniloquy is such a prevalent parasomnia that it is considered a normal sleep behavior among children.

Sleep Bruxism
Sleep bruxism was also found to be very common in early childhood. A longitudinal study reported that the prevalence of bruxism decreased progressively with age up to adolescence.9 This study shows that the prevalence increases at least between 2.5 to 6 years of age. In addition, the overall prevalence that was found here is much higher than what has been reported in previous studies of older children,21–23 which is consistent with the later progressive decrease in prevalence in adolescence.9 An age-related decline in prevalence also has been described throughout adulthood in a population-based study.24 Persistent sleep bruxism was found to be related to the parental behavior of staying with the child until he or she falls asleep. This could, in turn, be related to separation anxiety in the child, which also has been found to be associated to bruxism. Anxiety in general has been reported to be a factor that is associated with sleep bruxism in adolescents and adults.25,26

Sleep Terrors
The incidences that are reported for sleep terrors in the literature are wide-ranging.9,22,23,27 Discrepancies may be the result of the age range studied, the sampling method, and the definition used. Using an operational definition, we obtained a fairly high overall prevalence. This might reflect that our cohort of children was very young, probably at a period when sleep terrors are their peak of occurrence. However, it is nonetheless possible that some mothers mistook nightmares for sleep terrors. The observed association between sleep terrors and night wakings is probably attributable to the high level of behavioral and autonomic activity that accompanies sleep terrors. In addition, although a child can remain asleep through an entire sleep terror, even one with extensive episode of crying and yelling, the parents may perceive the child to be awake. The association that was found between somnambulism and sleep terrors, which has been reported before,8,9 is not surprising because these 2 parasomnias both arise from slow-wave sleep. Therefore, a predisposition to slow-wave sleep disturbance may lead to 1 or both parasomnias.

Separation anxiety emerged as a factor that was associated with both sleep terrors and somnambulism. Anxiety has been reported to increase the occurrence of parasomnias in adults and children, especially somnambulism and sleep terrors.26–30 On the basis of clinical and research experience, Rosen et al10 proposed that these 2 disorders of arousal might be the nocturnal expression of otherwise repressed anger feelings toward life events (eg, separation, divorce, parental conflict, family moves), which also translates into separation anxiety.

Somnambulism
The overall prevalence of somnambulism that we found for preschoolers (14.5%) is slightly lower than that reported for older children in other studies.28,31,32 However, in our cohort, the occurrence of new cases was still on the rise at the last measurement (age 6). Indeed, the peak of occurrence for somnambulism (~17%) has been described to be at ~11 to 12 years.32 Although many studies found no gender difference for somnambulism in older children and adolescents,9,32 it was recently reported to be more common in boys than in girls, a finding that is consistent with our results, in a large cohort of children who were aged 4 to 9 years.31 An explanation for the relationship between somnambulism and high hyperactivity-inattention scores is unclear. However, it has been reported that hyperactivity is increased in children with sleep problems13,34 or daytime sleepiness35 and that hyperactivity-inattention disorder often disappears when sleep disorders are treated.34,36 It remains to be clarified whether children with persistent somnambulism are more sleepy during the daytime, although our results indicate that their nocturnal sleep time is not affected.

Enuresis
Enuresis was assessed only at ages 5 and 6 because the clinical diagnosis of this condition can be made only for children who are 5 years and older. Nonetheless, the prevalence that was found here is similar to that reported by 2 population-based studies,17,38 which found that 20% to 33% of children were bed-wetting at the age of 5 years. The male predominance that was observed in this study for enuresis also has been well described.9,37–41 Our results suggest that enuresis is not linked with anxiety in preschoolers, although other studies have shown a relationship in older children.42–44 Anxiety probably emerges when the problem is long lasting. The child then is older, has more of a social life, and therefore is more self-conscious.

Rhythmic-Movement Disorder
The least prevalent parasomnia, rhythmic-movement disorder, was experienced by <10% of children during the study period. It is a parasomnia that both appears and disappears very early in life. The peak in incidence
has been determined to be between 9 and 18 months.\(^4^5\)
Indeed, the prevalence of rhythmic movements was reported to be higher (21%) in children of the same cohort studied at 17 months.\(^4^8\) Nonetheless, the annual frequency that was reported in this study is still lower than what was reported in other studies for the same age range.\(^5^5,4^6\) Sleep-related rhythmic-movement disorder was found to be associated with depression in the mother and low family SES. Body-rocking is usually considered a pleasant self-soothing behavior associated with sleep onset, but it has been suggested that it reflects an unbalanced mother–child relationship.\(^4^7\) More realistically, the probability that a child uses self-soothing rhythmic movements to fall asleep might be greater in the face of adversity (depression of the mother and stress in the family as a result of low SES) than in a more stable household.

Finally, it seems that the presence of childhood parasomnias, even persistent ones, has little or no impact on sleep duration. This is noteworthy because sleep duration affects various spheres of child development, such as cognition,\(^4^8,4^9\) physical development (BMI),\(^5^0,5^1\) and behavior.\(^5^2,5^3\)

**Limitations**
Some sleep problems that are known to affect children, such as sleep apnea syndrome, restless legs syndrome, and nightmares, were not reviewed in this article. Furthermore, comorbid health problems that could have an influence on dyssomnias or parasomnias were not studied. It must be said, however, that children with important medical problems were not included in this cohort. Before generalizing, then, one should keep in mind that the population studied here is mostly white and might not be representative of other populations. Objective sleep laboratory data also were not available to validate parental reports. However, a high correlation has been reported between the parents’ estimates of young children’s sleeping hours and their sleep time measured in the laboratory.\(^5^4\) Finally, our study design allowed the identification of factors that covary with the presence of dyssomnias or parasomnias, but it does not reveal causal relationships. Despite these limitations, however, the study design simultaneously evaluates various early childhood dyssomnias and parasomnias in a prospective and longitudinal manner and clarifies their associations with other domains of the child’s life.

**ACKNOWLEDGMENTS**
This study was funded by the Ministry of Health and Social Services (Quebec City, Quebec, Canada); the Canadian Institutes of Health Research (Ottawa, Ontario, Canada); the Social Sciences and Humanities Research Council of Canada (Ottawa, Ontario, Canada); the Quebec Fund for Research on Society and Culture (Quebec City, Quebec, Canada); the Quebec Fund for Research on Nature and Technology (Quebec City, Quebec, Canada); the Health Research Fund of Quebec (Montreal, Quebec, Canada); the Molson Foundation (Montreal, Quebec, Canada); the Ministry of Research, Science and Technology (Quebec City, Quebec, Canada); Human Resources Development Canada (Ottawa, Ontario, Canada); the Canadian Institute for Advanced Research (Toronto, Ontario, Canada); Health Canada (Ottawa, Ontario, Canada); the National Science Foundation (Arlington, VA); Montreal University (Montreal, Quebec, Canada); Laval University (Quebec City, Quebec, Canada); and McGill University (Montreal, Quebec, Canada). Drs Montplaisir and Tremblay hold Canada Research chairs on sleep disorders and child development, respectively.

**REFERENCES**
Impact of Medicaid Disenrollment on Health Care Use and Cost

Mary E. Rimsza, MD, Richard J. Butler, PhD, William G. Johnson, PhD

School of Computing and Informatics, Arizona State University, Tempe, Arizona; Department of Economics, Brigham Young University, Provo, Utah

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 health care use of children who are covered by public insurance and uninsured children who live in a large urban area and the potential impact of disenrollment on health care use and costs if these children become uninsured.

Methods. The 2004 health care transactions for 43,313 uninsured children and 168,722 children who were insured by Medicaid/State Children’s Health Insurance Program and living in the Phoenix metropolitan area were analyzed using a community-wide administrative health database (Arizona HealthQuery). Using a multivariate model of health care use by currently uninsured children, we examined the effect of 10% disenrollment of the children who were currently insured by Medicaid/State Children’s Health Insurance Program.

Results. A 10% disenrollment would increase the costs of health care in the community by $3,460,398 annually, or $2,121 for each child disenrolled. This increase in costs is attributed to a shift of care from ambulatory settings to more expensive emergency departments and an increase in hospital days. We determined that 69% of the change in emergency department visits, 58% of the change in hospital stays, and 74% of the change in ambulatory visits would be attributable to the change in insurance status.

Conclusions. Programmatic changes that result in disenrollment from public insurance programs will increase the number of emergency department visits and hospital days as well as the total community costs of health care. These increases in health care use can be expected to aggravate community problems of emergency department overcrowding and inpatient bed shortages. The majority of the changes in use are attributable to changes in insurance status, which results in a shift of care from less expensive ambulatory settings to emergency departments and increases in hospital days when children lose Medicaid/State Children’s Health Insurance Program coverage.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2747
doi:10.1542/peds.2006-2747

Key Words
Medicaid, uninsured children, emergency department, health care costs, SCHIP

Abbreviations
SCHIP—State Children’s Health Insurance Program
AZHQ—Arizona HealthQuery
ED—emergency department
CTS—Community Tracking Study

Accepted for publication Nov 8, 2006
Address correspondence to Mary E. Rimsza, MD, Center for Health Information and Research, School of Computing and Informatics, Ira A. Fulton School of Engineering, Arizona State University, PO Box 874711, Tempe, AZ 85287-4711. E-mail: mrimsza@asu.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
In an effort to contain costs, many states have made programmatic changes in their public health insurance programs (Medicaid/State Children’s Health Insurance Program [SCHIP]) that have resulted in disenrollment of children. A recent survey of state Medicaid program directors revealed that every state implemented at least 1 new Medicaid cost containment measure in fiscal year 2005.1 The most prevalent measures involved freezing provider payment rates, increasing premiums, or raising copayments. Although reducing the number of children who are enrolled in public insurance programs will, all else equal, reduce state Medicaid and SCHIP expenditures, the reductions in these expenditures do not represent an equal reduction in expenditures for the health care of the affected people from the viewpoint of state or federal governments or society as a whole. The net effect of these cost containment strategies can be understood only by considering all of the immediate and subsequent effects of the changes.

People who disenroll from Medicaid or SCHIP programs as a result of programmatic changes are not likely to replace public program insurance with private commercial insurance,2,3 because it is unlikely that people who live near the poverty level will be able to afford the premiums that are associated with private commercial health insurance. This report examines the effect of disenrollment on both health care use and costs by estimating the impact of a 10% decrease in Medicaid/SCHIP enrollment in Arizona. The assumed value of 10% represents a lower bound on the likely impact of most changes that are proposed by state governments. Most of the changes in cost-sharing or eligibility that were observed in other states have produced much larger reductions in enrollment, even for very small increases in costs, to people who are insured by Medicaid.4,5

METHODS

The 2004 health care transactions for children who were younger than 18 years and lived in Maricopa County, Arizona, which includes the Phoenix metropolitan area, were analyzed using a community health database, Arizona HealthQuery (AZHQ). This database currently includes health care transactions on >4 million Maricopa County residents, including all Medicaid and SCHIP participants and patients who receive care at federally qualified community health centers, the county’s public hospital and its affiliated clinics, a religiously affiliated free clinic, and most of the private hospitals in the county. Arizona’s SCHIP program is an expansion of its Medicaid program and administered by the same agency. Families apply for both SCHIP and Medicaid using the same application and are allocated to SCHIP or Medicaid on the basis of family size and income. Place of service was divided into 3 categories: emergency department (ED) visits, inpatient hospitalizations, and ambulatory care visits. Visits to private physician’s offices, community health centers, and hospital-based ambulatory care centers were considered ambulatory care visits.

The financial and health care use impact of a 10% disenrollment of the Medicaid/SCHIP–insured children was estimated, assuming that all of the disenrolled children became uninsured and these children would exhibit medical usage patterns just like the currently uninsured who had the same characteristics (age, gender, race, ethnicity, provider, and care for ambulatory care–sensitive conditions). An ambulatory care–sensitive condition was defined as a condition that could be treated in a primary care setting if timely care were provided and includes conditions such as gastroenteritis and urinary tract infections.7 We assumed that children who disenroll would become uninsured because only a small percentage of Medicaid/SCHIP-enrolled children have access to employer-sponsored coverage and the premiums for this employer-sponsored coverage are usually not affordable for families who are eligible for Medicaid/SCHIP.1 We combined the estimated coefficients of a multivariate model of health care use by currently uninsured children with children who are currently enrolled in Medicaid/SCHIP and then assumed that 10% of the latter group become uninsured. A system of logistic regression equations was used to estimate the probability of ED visits, inpatient days, and ambulatory office visits, and a system of 3 nonlinear regression equations was used to estimate the quantity of services for each of the 3 services if 10% of the Medicaid/SCHIP recipients become uninsured.

Each system of equations is estimated for all Medicaid/SCHIP-insured and uninsured children. Differences in use for uninsured and Medicaid/SCHIP-insured children are influenced by differences in demographic characteristics and by differences in insurance coverage. We used a regression decomposition, also known as the Oaxaca decomposition in the econometrics literature, to evaluate the differences between encounters by the uninsured and encounters by the Medicaid/SCHIP insured.8,9 The regression decomposition was modified to fit health care comparisons, to separate differences in use between Medicaid/SCHIP-insured and uninsured children into differences as a result of the characteristics of the people and differences as a result of insurance. This decomposition separates observed differences in outcomes (probabilities of use or quantities of use) into
differences that are associated with a person’s demographic variables and differences as a result of the insurance coverage.

The net changes in use for Medicaid/SCHIP-insured and uninsured children if 10% of the Medicaid/SCHIP-insured children became uninsured were calculated by multiplying the quantities of services by the numbers of children in each group before and after the simulated 10% disenrollment. Total health care expenditures are based on 2004 payments for ambulatory care visits, inpatient hospitalizations, and ED visits paid by Arizona’s Medicaid program. The average amount paid for each service was multiplied by the quantity of each service before and after the simulated disenrollment to estimate the aggregate change in health care expenditures. Because most of the health care costs for the uninsured are paid by federal and state governments through Medicare and Medicaid in the form of disproportionate share hospital adjustments and indirect medical education payments in addition to other federal programs, such as funding for community health centers and the Maternal and Child Health Bureau, we assume that the cost of care for children who become uninsured are community costs. Indeed, federal and state funds have been estimated to cover 87% of the total costs of uncompensated care. Potential savings from programmatic changes in Medicaid/SCHIP also are offset by increased Medicaid medically needy spending, increased tax subsidies to private insurance, and increased costs that are associated with uncompensated care.

The Oaxaca decomposition is a mathematical technique that first was used to measure discrimination. We used this technique to separate the difference in the dependent variable between the 2 insurance groups into the difference that was attributable to observable characteristics (ie, the portion of the difference that is explained by differences in the mean characteristics that are included in the model) and unobserved factors (ie, the portion of the difference that is attributable to differences in the coefficients between the 2 groups). The difference that is attributable to unobserved factors was considered to be the measure of the effect of insurance.

RESULTS
There were 40 945 uninsured children and 163 742 Medicaid/SCHIP-insured children in the AZHQ database for whom complete data were available on all demographic and health care use variables that were used in the analysis. The majority of the children in both insurance groups were from racial or ethnic minorities. The most common ethnic group was Hispanic: 62% of the Medicaid/SCHIP-insured children and 56% of the uninsured children were Hispanic (Fig 1). Children were divided into 4 age groups: 0 to 4 years, 5 to 9 years, 10 to 14 years, and 15 to 17 years. Within the 0- to 4-year age group, the costs were similar for 0- to 1-year-olds and the entire group. The majority of the Medicaid/SCHIP-insured children and uninsured children were younger than 10 years: 76% of the Medicaid/SCHIP-insured and 69% of the uninsured children (Fig 2). Fifty percent of the children were male in both the Medicaid/SCHIP-insured and uninsured groups. For the uninsured children, there were 800 ED visits per 1000 children in the year compared with 500 per 1000 children who were enrolled in Medicaid/SCHIP. The uninsured also had a higher average number of hospital days, averaging 1500 hospital days per 1000 children in the year compared with 500 hospital days for the Medicaid/SCHIP-enrolled children. The Medicaid/SCHIP-enrolled children averaged 2200 ambulatory visits per 1000 children in the year compared with 600 visits per 1000 children.

FIGURE 1
Racial and ethnic distribution of children.
for the uninsured. The total number of visits by facility for uninsured children and Medicaid/SCHIP-insured children is shown in Fig 3.

Figure 4 shows the annual changes in ED and inpatient service use if 10% (16,327) of the Medicaid/SCHIP-enrolled children disenroll and become uninsured. The total number of uninsured ED visits would increase from 32,076 to 43,716, and the total number of uninsured hospital days would increase from 60,570 to 78,251. The total number of ED visits in the community would increase by 2,772, and the total number of hospital days would increase by 9,273.

The total number of ambulatory visits by uninsured children would increase from 21,013 to 30,812 with a 10% disenrollment, and the number of ambulatory visits by Medicaid/SCHIP-insured children would decrease from 356,849 to 321,327. The net effect on service use in the ambulatory setting associated with a 10% disenrollment is that the total number of ambulatory visits would decrease by 25,723.

To determine the total financial effect of disenrollment on health care costs, we used the mean payments for uninsured visits by facility type from AZHQ data and multiplied the cost of the visit by the change in number of visits if 10% of the currently enrolled Arizona Health Care Cost Containment System patients were disenrolled and became uninsured. The mean payment for visits in 2004 was $795.57 for ED visits, $586.55 for inpatient hospital visits, and $162.64 for ambulatory visits. These amounts are paid by Arizona’s Medicaid program and include all services that are provided in these settings, including pathology, laboratory, radiology, and physician services. The payments are ~30% of the charges for these services. Assuming that these payments are indicative of the true cost of care, ED care would increase by $2,205,320 and the cost of inpatient

---

**FIGURE 2**
Age distribution of children versus insurance status.

**FIGURE 3**
Number of health care services according to facility, 2004.
care would increase by $5,439,078. However, the cost of ambulatory care would decrease by $4,183,589. The net increase in total costs would be $3,460,899, or $2,121 for each child who was disenrolled. Using the Oaxaca decomposition, we determined that 69% (8,031) of the ED visits, 58% (10,254) of the hospital days, and 74% (7,251) of the change in ambulatory visits would be attributable to the change in insurance status.

**DISCUSSION**

This study demonstrates the changes in health care use and costs if a randomly chosen 10% of the children who are currently enrolled in Arizona’s Medicaid/SCHIP program disenrolled and became uninsured. Our analysis of the health care encounters for both the Medicaid/SCHIP-insured and uninsured children is a function of each child’s age, gender, race, type of treating physician, ambulatory care–sensitive condition, and place of service. In fact, the overall characteristics of the uninsured and Medicaid/SCHIP-insured children is a function of each child’s age, gender, race, type of treating physician, ambulatory care–sensitive condition, and place of service. In fact, the overall characteristics of the uninsured and Medicaid/SCHIP-insured populations are so similar that when the likelihood of being uninsured is estimated as a function of these variables, the populations are virtually indistinguishable. Thus, the populations are not different because of observable characteristics but because they are equally likely to be uninsured given their characteristics. Because we hold observable characteristics and broad health indicators constant, it is likely that the differences observed are attributable to insurance coverage.

This analysis is not based on actual changes that have occurred in the Arizona Medicaid/SCHIP program or changes in other states but rather is a model for the effects of any programmatic change (eg, increased premiums; enrollment caps; decreased eligibility; enrollment barriers such as more frequent renewals, citizenship documentation) that results in disenrollment. Although we modeled the effect of a 10% disenrollment, most of the changes in cost-sharing or eligibility that have been enacted in other states have produced much more than a 10% disenrollment. For example, the outcomes of increased cost-sharing as a result of increases in premiums of 1% to 5% of family income by Washington’s Basic Health Plan, Minnesota’s Minnesota Care, and Hawaii’s Quest program resulted in reductions in enrollment that ranged from 18% to 57%, and Texas recently experienced a 29% reduction in SCHIP enrollment in <1 year after increasing premiums, adding a 90-day waiting period for benefits, and reducing the enrollment period from 12 months to 6 months.

Our simulation assumed that children who leave the Medicaid/SCHIP program would be random with respect to their medical needs. This is plausible if the policy change that results in disenrollment affects children with a random set of health needs, such as enrollment caps. Some policy changes that result in disenrollment may not be random. For example, increased paperwork requirements (eg, proof of citizenship) may be more likely to affect healthy children than chronically ill children because the parents of chronically ill children may be more willing to take the time to complete the paperwork requirements; in this case, our estimates would be an upper bound on the actual cost increases. However, if disenrollment is generated through a policy of higher copayments for hospitalizations, and ED and ambulatory visits, then this might disproportionately affect those with potentially the greatest medical needs and our disenrollment estimates would be a lower bound on the actual cost increases.

Our study demonstrates that the uninsured are more likely than Medicaid/SCHIP-insured children to receive health care in expensive sites (the ED and hospital) and less likely to receive care in the ambulatory setting. These findings are similar to our findings in Yuma, Arizona, where we found that uninsured children were nearly 4 times more likely to use the ED than insured children for ambulatory care–sensitive conditions. This increase in ED and hospital care may be attributable to the lack of preventive services and timely access to health care for acute care problems if a child is uninsured. Because the uninsured often are unable to obtain health care in physicians’ offices, they go to the ED, where they know they cannot be turned away because
of the Emergency Medical Treatment and Active Labor Act. However, the cost of nonemergent care that is provided in the ED is much higher than similar services that are provided in physicians’ offices.

Because of the shift in care from ambulatory sites to EDs and hospitals that is associated with disenrollment, our study demonstrates that the total number of ED visits and hospital days in the community increases as the number of children who disenroll increases. Our findings differ from Cunningham,1 who reported that Medicaid/SCHIP cuts that resulted in disenrollment would not affect the total number of ED visits. However, his analysis was based on 2000–2001 and 2003 Community Tracking Study (CTS) household surveys rather than administrative health data. In the CTS, the number of ED visits is self-reported, and those that resulted in hospitalization were excluded. Also, in the CTS survey, insurance coverage was determined at the time of the interview and not the time of the visit, as in our study. Finally, Cunningham assumed that 12% of the disenrolled children would acquire private insurance, whereas we assumed that these children would become uninsured.

Most states make programmatic changes in their Medicaid/SCHIP programs in an effort to decrease their costs for these programs. However, this study demonstrates that programmatic changes that result in disenrollment actually increase the total health care costs for the community. Most of the health care costs for the uninsured are paid by federal and state governments through Medicare and Medicaid in the form of disproportionate share hospital adjustments and indirect medical education payments in addition to other federal programs, such as funding for community health centers and the Maternal and Child Health Bureau. Indeed, federal and state funds have been estimated to cover 87% of the total costs of uncompensated care.11 Potential savings from programmatic changes in Medicaid/SCHIP also are offset by increased Medicaid medically needy spending, increased tax subsidies to private insurance, and increased costs that are associated with uncompensated care.2 In a previous analysis of Medicaid disenrollment in an agricultural community, we concluded that 10% disenrollment would increase the number of uninsured children by 21% and increase the community’s health care costs as a result of a shift in sites of care from less expensive ambulatory office sites to more expensive EDs and increased hospitalizations.12

The behavioral model of health care use and access by Anderson et al6 proposed that use is a function of need, enabling factors, and characteristics of the individual and his or her community. In this study, we used a community health database to assess the effects of Medicaid/SCHIP policy changes on health care use and costs. Following the model of Andersen et al,6 we included in our analysis personal characteristics that may affect use, including age, gender, race/ethnicity, and amount of care that is for ambulatory care–sensitive conditions. Although perceived and actual health needs are more difficult to assess, we did include in our analysis past visits for ambulatory care–sensitive conditions as well as specialty type of primary care provider. By using a community health database, community factors (eg, Medicaid/SCHIP acceptance by the community, availability of safety net providers, location of EDs) also are included in the analysis. Using the Oaxaca decomposition, our analysis demonstrates that the most important of these factors in determining health care use by facility is insurance status.

Recent changes in Medicaid requirements that requires US citizens to provide proof of citizenship for Medicaid enrollment may lead to increased disenrollment in Arizona. We hope to assess the accuracy of our simulation during the next 2 years using our AZHQ database to compare health care use of Arizona Medicaid/SCHIP recipients who disenroll because of lack of citizenship documents.

There are some limitations to our study. First, it is based on data from a large, urban community in the Southwest, where the majority of the Medicaid/SCHIP population as well as the uninsured are Hispanic. Second, the costs of disenrollment were based on Medicaid/SCHIP payments in Arizona. These payments are ~30% of charges for such services in Arizona. The actual costs are likely somewhere between these payments and charges. Therefore, our results likely underestimate the true costs of disenrollment. Third, although all ambulatory visits to private physicians’ offices for the Medicaid/SCHIP-insured children are included in the analysis as well as ambulatory visits for both the Medicaid/SCHIP-insured and uninsured patients to federally qualified community health centers, public clinics, and free clinics, some ambulatory visits for uninsured patients may not be included if they receive ambulatory care in other sites. This would result in an underestimation of the costs of disenrollment. Fourth, we assumed that children who disenrolled from Medicaid/SCHIP would become uninsured because either they were not eligible for private insurance or their families could not afford private insurance premiums. However, some low-income families may be able to gain access to private health insurance.

CONCLUSIONS

This study shows that cost containment strategies that result in disenrollment from Medicaid or SCHIP programs can be expected to increase care at expensive sites, including EDs and hospitals, and decrease care that is received in physicians’ offices. Such changes in site of care not only will increase health care costs but also will aggravate current community problems of ED overcrowding and inpatient bed shortages. The total number
of ED visits and inpatient days can be expected to increase as well as the amount of uncompensated care that is provided by hospitals and EDs that serve as safety net providers when Medicaid/SCHIP-insured children dis-enroll.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of St Lukes Health Initiatives and the cooperation of the Arizona Health Care Cost Containment System in the preparation of this report.

REFERENCES

Health Insurance Across Vulnerable Ages: Patterns and Disparities From Adolescence to the Early 30s

Sally H. Adams, RN, PhD, Paul W. Newacheck, DrPH, M. Jane Park, MPH, Claire D. Brindis, DrPH, Charles E. Irwin, Jr, MD

Department of Pediatrics, Institute for Health Policy Studies, and Center on Social Disparities in Health, University of California, San Francisco, California

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

ABSTRACT

OBJECTIVE. Young adults have the lowest rate of insurance coverage of any age group. Little is known about insurance patterns from adolescence through the early 30s. The objective of this study was to assess patterns and disparities in health insurance from adolescence through the early 30s.

DESIGN. We analyzed data from the 2002 and 2003 National Health Interview Survey (ages 13–32; N = 48,827). We examined public and private insurance coverage and conducted logistic regression to evaluate racial/ethnic and income disparities in coverage. Outcomes were insurance coverage at ages 13 to 32.

RESULTS. Insurance patterns follow a U-shaped curve across the age categories. Rates are highest at ages 13 to 14, lowest at ages 23 to 24, and then increase gradually. Private rate patterns are similar; however, public coverage decreases across ages. In bivariate analyses, black and Hispanic groups had lower coverage rates than the white group, and the low- and middle-income groups had lower rates than the high-income group. After adjustment for confounding variables, all disparities remained significant except for differences between the black and white groups.

CONCLUSIONS. After age 18, all groups are vulnerable to lack of insurance. Rate increases beyond age 25 to 26 years are attributable to increases in private coverage, whereas decreases in public coverage account for the lack of a full recovery to the higher rates seen in adolescence. The safety net of public programs that cover adolescents disappears in young adulthood, leaving young adults vulnerable, a problem that persists into the 30s for those who are in poverty and those who are of Hispanic origin.
Lack of health insurance is a critical issue during young adulthood. Young adults without health insurance, compared with their insured counterparts, are more likely to have unmet medical needs, to have no usual source of care, and to report fair or poor health. Young adults have the lowest insurance rate of any age group, with 32.5% of men and 26.9% of women aged 18 to 24 years uninsured in 1998–2000. Adults aged 19 to 29 years compose 17% of the population that is younger than 65 years but 30% of uninsured. Furthermore, 19- to 29-year-olds accounted for 40% of the growth in the uninsured population since 2000.

Recognizing the importance of continuous insurance coverage for this age group, the American Academy of Pediatrics and the Society for Adolescent Medicine issued policy/position papers emphasizing health insurance coverage and affordable health care for children, adolescents, and young adults. The Society for Adolescent Medicine also issued a position paper that spells out specific responsibilities for pediatric and other health care providers and addresses the transition to adult care for adolescents with chronic conditions. Among these is the elimination of policies, protocols, and restrictions of third-party payers that impede the timely transition to adult services for this population.

Changes in family and legal status that occur in the early 20s shape young adults’ ability to secure insurance. Coverage for adolescents has increased in the past 10 years, as declines in family-based employer coverage have been offset by increases in coverage under public programs, notably Medicaid and the State Children’s Health Insurance Program. In the transition into young adulthood, however, many young people lose eligibility under their parents’ private insurance coverage, and coverage generally ends at 21 years of age for the public programs that cover poor and near-poor adolescents. Those in the workforce are less likely to have employer-sponsored insurance, in part because young adults are more likely to be employed in low-wage positions and in settings with few employees.

Recent research has identified several factors that are associated with health insurance status among young adults. Women have slightly higher insurance rates than men. This may be because public programs are available to single mothers, although this has not been studied extensively. Rates of insurance coverage are lower among low-income and middle-income young adults. White individuals have the highest insurance coverage rates, followed by black and Hispanic individuals. Insurance rates are higher among young adults who are enrolled in school, especially those who are enrolled full time. In 2003, 39% of young adults who were aged 19 to 23 years and not enrolled in school full time were uninsured.

Traditionally, insurance is measured dichotomously, comparing those with any insurance at a point in time or during a specific period with those with no insurance. However, for many, insurance coverage may be intermittent or discontinuous. Therefore, recent research examined insurance status in a more comprehensive manner, by assessing coverage for a particular time frame, often 1 year, and taking into account intermittent insurance coverage as separate from having or lacking insurance for that time frame. A growing literature suggests that unstable or intermittent insurance also affects access to care. Children and adults with gaps in coverage experience substantially reduced access to care, compared with those with continuous coverage. Similarly, those who experience “insurance churning” (ie, losing and gaining coverage multiple times) are less likely to receive follow-up care and preventive services. Young adults have the highest rates of intermittent insurance and churning of any age group.

Most research on young adult health insurance either focuses exclusively on young adults or compares young adults with older age groups. Although these analyses have yielded important information, they have not examined how insurance status changes during the course of young adulthood. Although it is clear that adults who are older than 30 years have higher insurance rates than those who are in their 20s, we are not aware of any studies that have examined the dynamics of insurance during this period of the life span, as young people transition from adolescence into young adulthood and adults reach their late 20s and early 30s.

In this article we present a new analysis of patterns of health insurance, both public and privately funded, among young people from early adolescence through their early 30s. Our main objective was to describe the changes that take place in insurance coverage during this period of the life span. Our analysis demonstrates trends in private and public coverage, as well as full-year and part-year coverage. We also document differences in insurance patterns by socioeconomic status and among racial and ethnic groups during this period of transition.

Our analysis uses data from the 2002 and 2003 editions of the National Health Interview Survey (NHIS).

Methods

Sample and Population Characteristics

We used data from the 2002 and 2003 NHIS to conduct this study. The NHIS is a continuing nationwide household survey that is designed to collect information on the demographic characteristics, health status, and health care use patterns of the US civilian noninstitutionalized population. The survey has 3 modules: a basic module, a periodic module, and a topical module. The basic module contains 3 components: the family core, the sample adult core, and the sample child core. The family core includes a short set of questions that are administered for all family members, and the variables that were used for
the analyses herein were taken from the family core. The NHIS was selected for these analyses because it includes the pertinent variables of interest and has sufficient sample size to allow for analyses of narrow age groupings.

A total of 48,827 respondents who were aged 13 to 32 years were included in this study (24,559 from 2002 and 24,268 from 2003). A knowledgeable adult (typically a parent or guardian) answered questions about health insurance for adolescents who were younger than 17. When present, individuals who were 17 years and older answered the questions for themselves. The response rate for the NHIS family core questionnaires was 88.1% in 2002 and 87.9% in 2003.

**Variable Construction**

The family core includes information on sociodemographic characteristics, health insurance coverage, and family income. We constructed the health insurance variable on the basis of NHIS questions regarding health insurance coverage in the past 12 months. Respondents were asked whether they were currently insured. Furthermore, the NHIS asked those who were currently insured about the type of insurance that they had and whether they had been uninsured at any time in the past 12 months. Similarly, those who were currently uninsured were asked how many months it had been since they had been insured. From these NHIS responses, we constructed an insurance variable that consisted of 3 categories: (1) full-year insured, consisting of private and/or public insurance; (2) full-year uninsured; and (3) partial-year uninsured.

We examined health insurance status from ages 13 through 32 years, using 2-year age intervals (eg, 13–14, 15–16) to improve the statistical precision of our estimates. This 20-year age range was selected because it shows the insurance patterns of adolescence, young adulthood, and the years after young adulthood into the early 30s.

The family questionnaire provided information about family income and the ratio of annual family income for the past year to the federal poverty thresholds, which are adjusted for family size. For example, the poverty threshold for a family of 4 was $18,392 in 2002 and $18,810 in 2003. In this study, a poverty status index was used to create 3 income categories: below the poverty level and near poor (<200% federal poverty level [FPL]), moderate income (200%–399% FPL), and middle/high income (≥400% FPL). The nonresponse rate for income data, including missing and unknown responses, was 27%. Because >1 in 4 respondents did not report income, we used multiply imputed income values that were made available from the National Center for Health Statistics. We used SUDAAN statistical programs to analyze the multiply imputed data.

**Data Analysis and Presentation**

Estimates that are presented here were statistically weighted to reflect national population totals. The weights, provided by the data collection agency, are equal to the inverse of the sampling probability for each case, adjusted for nonresponse. Estimates, SEs, and test statistics were derived using SUDAAN software that takes into account the complex sample design of the survey, including household and intrafamilial clustering of sample observations.

The results are presented in graphic and tabular form. All analyses were conducted separately for each of the 2-year age groups, and all results are presented by age group. First, we present rates of insurance status in 3 categories: full-year insured, full-year uninsured, and partial-year uninsured. Second, we present rates for the full-year insured category separately according to whether the source of coverage was private or public. Third, we present rates of full-year insurance for white, black, and Hispanic groups. We conducted a bivariate logistic regression by using a dichotomous insurance outcome variable for each age group (full-year insured versus full- or partial-year uninsured), with race/ethnicity as the independent variable. This yielded unadjusted odds ratios (ORs) for the white, black, and Hispanic groups on the insurance variable. We also conducted a multivariate logistic regression for each age group that yielded adjusted ORs (aORs) for the racial/ethnic groups, controlling for income group, gender, and region of residence as independent variables. Fourth, we present rates of full-year insurance by the 3 income levels. We conducted a bivariate logistic regression using the dichotomous insurance outcome variable for each age group, with income level as the independent variable. This yielded unadjusted ORs for the income groups on the dichotomous insurance outcome. Using results of the multivariate logistic regressions described, we report the aORs for the income groups, controlling for race/ethnicity, gender, and region of residence.

**RESULTS**

**Health Insurance Coverage Across Age Categories**

Insurance coverage follows a U-shaped curve across the age categories (Fig 1). Rates of full-year insurance are highest among adolescents aged 13 to 14 years (87%) and lowest among 23- to 24-year-olds (61%). Rates then increase to 75% in the 31- to 32-year-old group, 12 percentage points lower than the rate for the 13- to 14-year-old group. The corresponding inverted U-curves indicate that full-year uninsurance rates are highest at ages 21 to 22 years and partial-year uninsurance rates are highest at ages 23 to 24 years. Taken together, the highest rate of any uninsurance occurs at ages 23 to 24 years.
Public and Private Health Insurance Coverage

Rates of private insurance coverage follow a U-shaped curve similar to that of the overall full-year coverage (Fig 2). Private coverage rates are stable from 13 to 18 years of age; they then decline to their lowest rates at ages 21 to 22 years and then increase to reach their highest levels at ages 31 to 32 years. Coverage rates for public insurance are highest at ages 13 to 14 years; they then decline most sharply between ages 17 to 18 and 19 to 20 years and then continue to decline gradually to a low of 7% by ages 31 to 32 years. The loss of public insurance throughout the late adolescent and young adult years accounts for the lack of a complete restoration of overall insured rates by ages 31 to 32 years to equal those at ages 13 to 14 years (Fig 1).

Racial/Ethnic Disparities in Health Insurance Coverage

Patterns of health insurance within the different racial/ethnic groups tend to follow a similar U-shaped curve across the age range (Table 1). The white group has the highest insured rates, and the Hispanic group has the lowest rates at each age. The gap between the white and the Hispanic groups increases from 16 percentage points at ages 13 to 14 years to a maximum of a 32–percentage point difference at ages 19 to 20 years and then fluctuates between 25 and 31 percentage points across the remaining age groups. The rate differences are statistically significant within each 2-year age interval. Although the differences between the white and Hispanic rates were somewhat attenuated when gender, income, and region of residence were controlled using multivariate analysis, they remained significant at each age.

Overall, our analyses found few significant differences between the white and black groups. When confounding variables of gender, poverty level, and region of residence are controlled, the rate differences remain significant in only 2 of the groups: ages 19 to 20 and 21 to 22 years. This suggests that the majority of the difference between the white and the black groups is accounted for by the confounding factors. The gap between the white and black groups is 3 percentage points at ages 13 to 14 years, widens during the early young adult years, and then diminishes to a 4–percentage point gap at ages 31 to 32 years.

Income Disparities in Health Insurance Coverage

Table 2 presents the rates of insurance for the 3 income groups and illustrates that the highest and middle-in-
come group rates follow a U-shaped pattern most clearly. The lowest income group shows the drop in coverage from ages 13 to 14 years through ages 25 to 26 years and then demonstrates very little recovery of rates after that. The gap between the lowest and highest income groups grows from 19 percentage points at ages 13 to 14 years to a maximum of 39 percentage points at ages 27 to 28 years and remains throughout subsequent ages. The insurance rate differences between the highest and both the middle and lowest income groups are significant at each age. When confounding factors of gender, race/ethnicity, and region of residence are controlled, the differences remain significant and show very little attenuation. This suggests that income has a strong independent effect on rates of insurance coverage.

DISCUSSION

In this article we have described the insurance patterns of Americans from adolescence through the early 30s. The analyses update and expand previous research that addressed insurance for adolescents and young adults. This study makes several unique contributions regarding health insurance patterns. First, showing changes across 2-year age increments allows us not only to highlight differences between the young adult years and adolescence and the early 30s but also to illustrate differences within the young adult years. Second, the presentation of insurance rates broken into full-year coverage and both partial- and full-year uninsurance allows us to present estimates and ORs that differentiate between those with or without stable coverage. This is important because previous research has shown that those with partial-year uninsurance were more similar to full-year uninsured groups in terms of decreased access to care9,11,13 and because a broad goal is continuous insurance for all. Finally, we examine insurance patterns for an expanded age range, spanning the major transitions from adolescence through the early 30s. Previous articles have documented coverage rates for adolescents or young adults (typically ages 18 through 24 or 25 years).

TABLE 1

<table>
<thead>
<tr>
<th>Age, y</th>
<th>Total Sample, %</th>
<th>White (Referent Group), %</th>
<th>Black</th>
<th>Hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>%</td>
<td>% OR (95% CI)</td>
</tr>
<tr>
<td>13–14</td>
<td>86.5</td>
<td>89.5</td>
<td>86.9</td>
<td>1.3 (1.0–1.7)</td>
</tr>
<tr>
<td>15–16</td>
<td>86.4</td>
<td>90.8</td>
<td>84.9</td>
<td>1.8 (1.4–2.2)</td>
</tr>
<tr>
<td>17–18</td>
<td>83.1</td>
<td>88.0</td>
<td>82.0</td>
<td>1.6 (1.3–2.1)</td>
</tr>
<tr>
<td>19–20</td>
<td>66.7</td>
<td>74.0</td>
<td>59.8</td>
<td>1.9 (1.6–2.3)</td>
</tr>
<tr>
<td>21–22</td>
<td>61.8</td>
<td>68.1</td>
<td>58.1</td>
<td>1.5 (1.3–1.9)</td>
</tr>
<tr>
<td>23–34</td>
<td>61.0</td>
<td>66.4</td>
<td>61.2</td>
<td>1.3 (1.0–1.5)</td>
</tr>
<tr>
<td>25–26</td>
<td>64.0</td>
<td>69.5</td>
<td>63.6</td>
<td>1.3 (1.1–1.6)</td>
</tr>
<tr>
<td>27–28</td>
<td>68.3</td>
<td>74.0</td>
<td>66.7</td>
<td>1.4 (1.1–1.8)</td>
</tr>
<tr>
<td>29–30</td>
<td>72.4</td>
<td>78.6</td>
<td>73.4</td>
<td>1.3 (1.1–1.7)</td>
</tr>
<tr>
<td>31–32</td>
<td>75.0</td>
<td>79.8</td>
<td>76.4</td>
<td>1.2 (1.0–1.6)</td>
</tr>
</tbody>
</table>

CI indicates confidence interval.
* Significant at P < .05.

TABLE 2

<table>
<thead>
<tr>
<th>Age, y</th>
<th>Total Sample, %</th>
<th>$&lt;$400% (Referent Group), %</th>
<th>$\geq$400%</th>
<th>200%–399%</th>
<th>FPL</th>
<th>%</th>
<th>% OR (95% CI)</th>
<th>aOR (95% CI)</th>
<th>%</th>
<th>% OR (95% CI)</th>
<th>aOR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13–14</td>
<td>86.5</td>
<td>95.9</td>
<td>88.6</td>
<td>3.0 (1.9–4.8)</td>
<td>1.3 (1.0–1.7)</td>
<td>2.8 (1.8–4.5)</td>
<td>77.0</td>
<td>7.0 (4.4–11.3)</td>
<td>6.1 (3.7–10.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15–16</td>
<td>86.4</td>
<td>96.7</td>
<td>87.8</td>
<td>4.0 (2.7–6.0)</td>
<td>1.3 (1.0–1.7)</td>
<td>3.5 (2.4–5.3)</td>
<td>75.6</td>
<td>9.4 (6.7–13.2)</td>
<td>7.1 (5.0–10.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17–18</td>
<td>83.1</td>
<td>95.0</td>
<td>86.3</td>
<td>3.0 (2.3–4.1)</td>
<td>1.3 (1.0–1.7)</td>
<td>2.8 (2.1–3.8)</td>
<td>70.4</td>
<td>8.0 (6.0–10.7)</td>
<td>6.5 (4.8–8.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19–20</td>
<td>66.7</td>
<td>82.6</td>
<td>66.0</td>
<td>2.4 (2.0–3.1)</td>
<td>1.3 (1.0–1.7)</td>
<td>2.2 (1.7–2.7)</td>
<td>58.0</td>
<td>3.4 (2.7–4.4)</td>
<td>3.0 (2.3–3.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21–22</td>
<td>61.8</td>
<td>75.0</td>
<td>62.1</td>
<td>1.8 (1.5–2.3)</td>
<td>1.3 (1.0–1.7)</td>
<td>1.7 (1.4–2.1)</td>
<td>54.9</td>
<td>2.5 (1.9–3.1)</td>
<td>2.2 (1.7–2.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23–34</td>
<td>61.0</td>
<td>74.3</td>
<td>62.5</td>
<td>1.7 (1.4–2.1)</td>
<td>1.3 (1.0–1.7)</td>
<td>1.7 (1.4–2.1)</td>
<td>49.7</td>
<td>2.9 (2.4–3.6)</td>
<td>2.8 (2.2–3.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25–26</td>
<td>64.0</td>
<td>79.4</td>
<td>67.6</td>
<td>1.9 (1.5–2.3)</td>
<td>1.3 (1.0–1.7)</td>
<td>1.7 (1.4–2.1)</td>
<td>46.1</td>
<td>4.5 (3.7–5.5)</td>
<td>4.0 (3.2–4.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27–28</td>
<td>68.3</td>
<td>85.9</td>
<td>70.9</td>
<td>2.5 (2.0–3.2)</td>
<td>1.3 (1.0–1.7)</td>
<td>2.3 (1.8–2.9)</td>
<td>46.8</td>
<td>6.9 (5.6–8.5)</td>
<td>6.1 (4.9–7.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29–30</td>
<td>72.4</td>
<td>88.5</td>
<td>73.1</td>
<td>2.8 (2.2–3.6)</td>
<td>1.3 (1.0–1.7)</td>
<td>2.6 (2.0–3.4)</td>
<td>51.0</td>
<td>7.4 (5.9–9.3)</td>
<td>6.4 (5.0–8.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31–32</td>
<td>75.0</td>
<td>90.1</td>
<td>77.7</td>
<td>2.6 (2.1–3.3)</td>
<td>1.3 (1.0–1.7)</td>
<td>2.6 (2.0–3.3)</td>
<td>51.1</td>
<td>8.7 (7.1–10.7)</td>
<td>8.2 (6.6–10.1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI indicates confidence interval.
* Significant at P < .05.
but no analysis has followed rates past young adulthood through the early 30s.

Our results show a steep decline in insurance rates after 18 years of age that continues through the mid-20s. More than 10.7 million young adults who were aged 19 to 26 were uninsured for all or part of the previous year. This was followed by a gradual increase in coverage that continued into the 30s. Rates in the early 30s do not recover to their earlier adolescent levels. The weak recovery in the late 20s and early 30s is directly related to the loss of public coverage, which diminishes steadily from adolescence onward.

Our analyses show that racial and ethnic disparities in coverage are present during adolescence and persist into young adulthood and the early 30s. Whereas disparities that are present in the black group are largely accounted for by income level, those that are observed in the Hispanic group, relative to the white group, remain significant when income level is controlled, indicating that additional factors account for some of the disparities in the Hispanic group. Income disparities in coverage continue to persist when we control for confounding factors, highlighting the importance of income in having coverage.

Several factors may account for the drop in coverage that was noted after age 18, regardless of income level. Although some employer-based insurance programs now provide coverage to dependents beyond age 18 or 19 years of age, regardless of student status, the majority of programs still do not provide extended coverage for nonstudents. In addition, public programs cover adolescents only until age 18 or 21 years, and for many who use the public programs, there are no affordable alternative sources for coverage. Our finding that public coverage decreases steadily after adolescence supports this. We also found that the largest drop-off in coverage occurred for traditionally disadvantaged groups, including Hispanic individuals and those with low incomes. Individuals in these groups are more likely to be employed in low-wage, part-time or temporary jobs or in jobs in small businesses that either do not offer insurance coverage or offer coverage that includes employee premiums that are unaffordable for many. Our analyses suggest that for those groups, the gaps in coverage are likely to extend beyond young adulthood, into the 30s.

In the past few years, policy makers at the federal and local levels have initiated programs and legislation to expand coverage to the young adult population. As many as 9 states passed or considered legislation to increase the age limit (ranging from 24 to 30 years) for dependent coverage on family policies. However, the effect of state insurance regulations may be limited because many large employers choose to self-insure and are exempt from state regulation. Several private insurers offer coverage for young adults now, either through individual or family plans. Several states mandate that colleges and universities require that students have health insurance and in some cases require that institutions make coverage available.

Although insurance gaps for young adults are being addressed at many levels, we lack an organized approach at the national level to address the problem. The drop-off in public coverage that we showed across each age group primarily affects lower income groups and those with special health care needs. Three regulatory changes would help to increase coverage: (1) increases in the age limit for State Children’s Health Insurance Program coverage; (2) provision of incentives for affordable employer-sponsored programs in low-wage job settings; and (3) changes in Medicaid eligibility criteria so that young adults with special health care needs qualify for coverage with broader allowances for the ability to work.

Pediatric health care providers can work at an individual level to help families and patients strategize options to maintain coverage as teens finish high school and plan future activities and can work at the policy level to publicize the importance of insurance coverage for all. Providers can help families and patients with special health conditions to plan and to minimize insurance gaps and can work at the policy level for the redefinition of criteria for young adults with health conditions to increase eligibility.

The primary limitation of this analysis is that the data are cross-sectional, not longitudinal. As a consequence, our analysis provides provisional information about how insurance status may change from adolescence through the early adult years. To be definitive, these findings need to be replicated using a longitudinal design.

Our analysis shows that lack of insurance is a problem for higher income young adults, as well the poor and near poor. We also show that the problem of lack of coverage persists into the late 20s, with the period between ages 19 and 28 years particularly at risk, and continues even into the 30s for those in poverty and those of Hispanic origin. We argue that in addition to increasing the age for dependents who are covered in private programs, increasing coverage for all must include improvements in publicly funded programs and improvements in coverage that is offered in low-wage employment settings. Improving coverage will result in measurable improvements in health and productivity during this transitional period for America’s young adults.

ACKNOWLEDGMENTS
This study was supported by Federal Maternal and Child Health Bureau grants U45 MC00023 and U45 MC00002.

We acknowledge the help of Michael Berlin, Tina Paul Mulye, Elizabeth Valitchka, and Jennifer Yu in the preparation of the manuscript.
REFERENCES

21. The Library of Congress. Health Care for Young Adults Act of 2005 (Bill No. H.R. 3040) [Bill introduced in Congress]
Psychological Functioning and Coping Among Mothers of Children With Autism: A Population-Based Study

Guillermo Montes, PhD, Jill S. Halterman, MD, MPH

OBJECTIVES. Studies suggest that having a child with autism has a negative impact on maternal psychological functioning, but no large-scale, population-based studies are available. The objectives of this study were to (1) describe the psychological functioning, physical and mental health, family communication, and parenting support of mothers of a child with autism compared with other mothers on a population basis and (2) assess the independent relationship between having a child with autism and these outcomes, controlling for the child’s social skills and demographic background.

METHODS. Mothers of 61,772 children who were 4 to 17 years of age were surveyed by the National Survey of Children’s Health, 2003. Autism was measured from an affirmative maternal response to the question, “Has a doctor or health professional ever told you your child has autism?” There were 364 children with autism in the sample.

RESULTS. Mothers of a child with autism were highly stressed and more likely to report poor or fair mental health than mothers in the general population, even after adjustment for the child’s social skills and demographic background. However, mothers of a child with autism were more likely to report a close relationship and better coping with parenting tasks and less likely to report being angry with their child after adjustment for the child’s social skills and demographic background. Having a child with autism was not associated with lower social support for parenting, an altered manner in which serious disagreements were discussed in the household, or increased violence in the household.

CONCLUSION. Mothers of children with autism showed remarkable strengths in the parent–child relationship, social support, and stability of the household in the context of high stress and poorer mental health.
Several studies recently considered the impact of having a child with an autism spectrum disorder (ASD) on parental psychological functioning.1–5 Uniformly, the literature finds that parenting a child with ASD is associated with higher levels of stress.1,6–8 Comparing parents of children with autism with parents of children without autism, one study found that parents of children with autism had much higher levels of family strain and lower social support.9 Another recent study2 showed that having a child with ASD was the primary contributor to maternal stress, even in the presence of other maternal mental health problems. Perceived social support has been identified as an important stress buffer that influences the parenting of mothers with a child with autism.9

Having a child with autism also influences marital and family functioning. Parents of children with autism are more likely to use avoidant coping strategies that may have a negative impact on the marital relationship.6,7 Several studies reported lower marital satisfaction, and at least 1 study10 showed that mothers of children with autism had lower marital satisfaction than mothers of children with Down syndrome.

Research on the impact of having a child with autism on the probability of divorce is inconclusive.1,11,12 Although not focused on psychological functioning, one population-based study13 found that children with autism in the United States were equally represented in 2-parent families compared with children without autism.

The main weakness of the previous literature on psychological functioning in parents of a child with autism is the potential for sampling bias. Many studies have drawn from clinic- or referral-based samples or autism societies for their data collection. As previous studies have noted,1,7 parents who belong to autism societies, are known to local organizations that serve children with ASD, or received care in specialty clinics likely are not representative of the general population of parents of children with autism. Because there are no population studies, it is not clear whether mothers of a child with autism perceive that they have lower parenting support than other parents, whether they are psychologically distanced from their child, or whether they rate their parenting as poor on a population basis. Moreover, little is known about which aspects of autism have a negative impact on maternal psychological functioning. One study14 found that the child’s behavior problems were more predictive of maternal stress than the severity of the child’s autistic symptoms. Other studies suggested that the stress of caring for a child with autism cannot fully explain psychological impairments in the mothers of children with autism.15 No studies have attempted to compare mothers of children with autism with mothers of other children with similar levels of social skills but who do not have a diagnosis of autism. We hypothesized that compared with mothers of other children with similarly poor social skills, the psychological functioning of mothers of children with autism will not be significantly impaired.

Recently, researchers used nationally representative telephone surveys to estimate the prevalence of autism in the United States on the basis of parental responses to a single question. All studies reported prevalence estimates within the generally accepted range of 6 per 1000, thereby supporting the validity of parent-reported information about autism.1,16 Using the National Survey of Children’s Health, 2003 (NSCH), we conducted the first population-based study of psychological and family functioning of mothers of children with autism. Our objectives were to (1) describe the psychological functioning, physical and mental health, family communication, and parenting support of mothers of a child with autism compared with other mothers on a population basis and (2) assess the independent relationship between having a child with autism and these outcomes, controlling for the child’s social skills and demographic background.

METHODS

Sample

The NSCH is part of the State and Local Area Integrated Telephone Survey program conducted by the National Center for Health Statistics. The survey interviewed 102 353 parents of children who were aged 0 to 17 between January 2003 and July 2004 (87% of the interviews were completed in 2003). The survey interviewed the adult in the household who was most knowledgeable about the sampled child’s health, typically the mother. The response rate was 68.8%. Weights were provided to adjust for the complex survey design and for nonresponse rates. Therefore, the estimates that are presented in this article are nationally representative estimates.

Analytic Sample

We restricted the sample to 61 772 school-aged children (4–17 years) for whom the respondent was the mother of the child sampled.

Measures

Autism Measure

Autism was measured with the following yes/no question: “Has a doctor or health professional ever told you that [child] has any of the following conditions? Autism?”

Physical and Mental Health

Maternal overall health was measured by the question, “Would you say that in general your health is excellent, very good, good, fair, or poor?” Mothers who responded
fair or poor were coded as having poor health. Similarly, mothers who responded fair or poor to the question, “Would you say that in general your mental and emotional health is excellent, very good, good, fair, or poor?” were coded as having poor mental health.

Parenting Stress and Parent–Child Relationship
We used items from the parent aggravation scale used in the National Survey of America’s Families studies to measure parent stress and aspects of the parent–child relationship. The scale was derived from the Parental Stress Index and Parental Attitudes About Childrearing scale. It contains 4 items: (1) how often in the last month the parent believed the child was much harder to care for than most, (2) how often the child did things that really bothered the parent, (3) how often the parent was giving up more of his or her life to meet the child’s needs than expected, and (4) how often the parent felt angry with the child. Items 1 and 3 are regarded as maternal stress items, whereas items 2 and 4 are categorized as parent–child relationship items in this study. All items were dichotomized as “never” versus “sometimes,” “usually,” and “always.” Closeness of the parent–child relationship was measured by a dichotomous variable that grouped parents who reported “very close” relationships versus those who reported relationships as “somewhat close,” “not very close,” and “not close at all.”

Coping With Parenting
Coping was measured by the following item: “In general, how well do you feel you are coping with the day-to-day demands of parenthood? Would you say that you are coping very well, somewhat well, not very well, or not well at all?” Parents who answered “very well” were coded as coping very well, whereas all other answers were coded as not coping very well.

Parent Support
Parent support was measured by a yes/no question that asked, “Is there someone that you can turn to for day-to-day emotional help with parenting?”

Family Communication and Domestic Violence
Two questions that were adopted from the National Survey of Families and Households and the Early Childhood Longitudinal Survey measured increasingly violent approaches to domestic disagreement. Parents were asked, “When you have serious disagreements with your household members, how often do you end up hitting or throwing things?” Those who responded with “never” were coded as having no domestic violence. Those who responded with any answer other than never were coded as having domestic violence, following the usage of the question in the literature on marital violence. Similarly, those who responded with any answer other than never to 2 similar questions (“When you have serious disagreements with your household members, how often do you argue heatedly or shout?” and, “How often do you just keep your opinions to yourself?”) were coded as having “heated arguments” and “keeping opinions to oneself,” respectively.

For the remaining disagreement style question, “When you have serious disagreements with your household members, how often do you discuss your disagreements calmly?” a different coding approach was used, because only 2 mothers of a child with autism responded with “never.” Therefore, we coded a mother who responded with “usually” or “always” as “discusses calmly.”

Child Prosocial Skills
We measured child prosocial skills by the sum of 4 items, some of which were adapted from the Positive Behaviors Scale. These items comprise a scale that is currently being validated by the Centers for Disease Control and Prevention. The 4 items were (1) “[he/she] shows respect for teachers and neighbors”; (2) “[he/she] gets along well with other children”; (3) “[he/she] tries to understand other people’s feelings”; and (4) “[he/she] tries to resolve conflicts with classmates, family, or friends.” Each item was scored on a 4-point scale (1 = never to 4 = always). Cronbach’s α reliability in the analytic sample was .649. The α reliability for the autistic subsample was .650.

Demographic Variables
Family structure was measured with a dichotomous 2-parent family indicator (2-parent households with biological, adoptive, or step parents versus single-parent families and all other family structures). Family poverty was coded using the derived NSCH variable that used income and family size information to classify the household’s income as above or below 200% of the 2003 federal poverty level, based on Department of Health and Human Services guidelines. Age, gender, race, and Hispanic ethnicity were measured by using direct questions from the survey. All participants consented to the study. Additional information on the survey and informed consent procedures are available.

Statistical Analysis
Stata (Stata Corp, College Station, TX) was used to adjust for the complex sample design using Taylor approximations that provide the correct standard errors, following NSCH guidelines. Demographic characteristics were analyzed using univariate descriptive and χ² statistics. We conducted multivariate logistic regressions on the outcomes after controlling for demographic variables (age, gender, race, Hispanic ethnicity, education, and family structure) and the child’s prosocial skill score. Because children with autism have a range of social
skills, the prosocial skill score was entered as a continuous variable.

RESULTS

Description of the Sample

Table 1 shows the demographic characteristics for the samples of children with autism and children without autism. There were 364 children with autism in the analytic sample. As previously reported, the prevalence rate for autism in the United States on the basis of the NSCH 2003 sample is 6 per 1000 (95% confidence interval [CI]: 4.7–6.8) with male individuals having a rate of 9 per 1000 (95% CI: 7.2–11.2) and female individuals a rate of 2 per 1000 (95% CI: 1.5–2.8). Parental report of autism probably includes diagnoses throughout the autism spectrum. Aside from a greater prevalence of autism among boys, children with autism were equally represented in the US population by race of the child, Hispanic ethnicity (any race), household education, and household poverty. There were no statistical differences between families with and without a child with autism by family structure (2-parent families). The average age of the children with autism (10.07 years; 95% CI: 9.50–10.64) was comparable to the average age of the children without autism (10.44 years; 95% CI: 10.38–10.49).

Parent–Child Relationship and Psychological Functioning: Univariate Results

Table 2 compares mothers of a child with autism with other mothers on the maternal health and psychological functioning variables. Mothers of a child with autism rated their overall health as poor or fair at similar rates as the general population. However, mothers with a child with autism were more likely to report poor or fair mental and emotional health (17.3% vs 7.1%; P < .001). A significantly higher percentage of mothers with a child with autism reported that their child was much harder to care for than most children his or her age compared with mothers without a child with autism (85.6% vs 30.8%; P < .001). Similarly, more mothers of children with autism reported that they had given up more of their life than expected during the previous month to meet the child’s needs (68.3% vs 40.9%; P < .001).

Of a child with autism, mothers with a child with autism were indistinguishable from other mothers (47.0% vs 79.3%; P > .05). Conversely, mothers of a child with autism reported that they were not able to share ideas or talk about the things that really mattered with the child as well as other mothers (47.0% vs 76.8%; P < .001). They also reported that their child “bothered them a lot” more frequently than other mothers (80.6% vs 64.8%; P < .001). On coping with parenting and parenting support, mothers with a child with autism were indistinguishable from other mothers. Approximately half of the mothers reported coping well with the demands of parenting (51.1% vs 54.3%; P > .05), and almost 9 of 10 reported having someone they “could turn to for day-to-day emotional help with parenting” (89.9% vs 86.3%; P > .05). Last, there were no detectable differences in the manner

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Demographic Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characteristic</strong></td>
<td>Family Has a Child With Autism</td>
</tr>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td>All</td>
<td>364</td>
</tr>
<tr>
<td>Female</td>
<td>79</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>294</td>
</tr>
<tr>
<td>Black</td>
<td>31</td>
</tr>
<tr>
<td>Multiracial</td>
<td>14</td>
</tr>
<tr>
<td>Other</td>
<td>10</td>
</tr>
<tr>
<td>Hispanic (any race)</td>
<td>28</td>
</tr>
<tr>
<td>High school education or less</td>
<td>74</td>
</tr>
<tr>
<td>Below federal poverty level</td>
<td>108</td>
</tr>
<tr>
<td>2-parent family</td>
<td>259</td>
</tr>
</tbody>
</table>

aP < .001.

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Bivariate Comparisons of Maternal Functioning Among Families With and Without a Child With Autism</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter</strong></td>
<td><strong>Family Has a Child With Autism</strong></td>
</tr>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td>Maternal health</td>
<td></td>
</tr>
<tr>
<td>Overall health (poor/fair)</td>
<td>47</td>
</tr>
<tr>
<td>Mental and emotional health (poor/fair)</td>
<td>51</td>
</tr>
<tr>
<td>Maternal stress</td>
<td></td>
</tr>
<tr>
<td>Child is harder to care for than most children</td>
<td>301</td>
</tr>
<tr>
<td>Parent gives up her life to meet child’s needs</td>
<td>246</td>
</tr>
<tr>
<td>Parent–child relationship</td>
<td></td>
</tr>
<tr>
<td>Relationship is very close</td>
<td>288</td>
</tr>
<tr>
<td>Talk about important things with child</td>
<td>157</td>
</tr>
<tr>
<td>Angry at child</td>
<td>279</td>
</tr>
<tr>
<td>Child bothers mother a lot</td>
<td>295</td>
</tr>
<tr>
<td>Coping and support</td>
<td></td>
</tr>
<tr>
<td>Coping very well with parenting</td>
<td>163</td>
</tr>
<tr>
<td>Has someone to talk about parenting</td>
<td>327</td>
</tr>
<tr>
<td>Disagreement style</td>
<td></td>
</tr>
<tr>
<td>Keeps opinions to oneself</td>
<td>259</td>
</tr>
<tr>
<td>Discusses calmly</td>
<td>260</td>
</tr>
<tr>
<td>Argues heatedly</td>
<td>271</td>
</tr>
<tr>
<td>Hits or throws things</td>
<td>43</td>
</tr>
</tbody>
</table>

aP < .01.
bP < .001.
in which serious disagreements were discussed in the household.

Parent–Child Relationship and Psychological Functioning: Multivariate Results
Table 3 shows the unadjusted and adjusted odds ratios (ORs) and 95% CIs of the association of having a child with autism on parent–child relationship and psychological functioning outcomes. Adjusted ORs resulted from models that controlled for demographic variables (age, gender, race, Hispanic ethnicity, household education, poverty, and family structure) and the child’s prosocial skills scale score. We specifically adjusted for the child’s social skills to determine whether maternal psychological functioning was different for mothers of children with autism compared with mothers of children with similarly low social skills.

As expected, the majority (61%) of children with autism scored ≥1 SD below the mean in the child’s prosocial skill scale. Children with low social skills and no diagnosis of autism were a large group (8450) of whom the majority (72%) had no behavioral diagnoses (8% had attention-deficit/hyperactivity disorder, 8% had behavior or conduct problems, and 11% had both).

In the adjusted model, mothers of a child with autism were more than twice as likely as mothers in the general population to report poor or fair mental and emotional health (OR: 2.42; 95% CI: 1.31–4.45). Similarly, more mothers of children with autism reported that they were giving up their life to care for their child beyond their expectations (OR: 3.77; 95% CI: 2.43–5.87).

Despite these challenges to their parenting and mental health, mothers with a child with autism showed some remarkable strengths. After adjustment for demographic background and the child’s prosocial skills, mothers of a child with autism were 5.2 times more likely to report a close relationship with their child (OR: 5.20; 95% CI: 2.05–13.20). In addition, mothers of a child with autism were 1.8 times less likely to report that they were angry at their child (OR: 0.54; 95% CI: 0.34–0.85).

Even after adjustment for demographic background and social skills, mothers of a child with autism reported that they were less likely to be able to talk to their child with autism about things that really matter (OR: 0.44; 95% CI: 0.29–0.67). It is interesting that mothers with a child with autism were no longer more likely than other mothers to report that their child bothers them more frequently (OR: 1.18; 95% CI: 0.73–1.93) in the adjusted analysis.

In terms of coping and support, mothers with a child with autism were more likely to report that they were coping well with the day-to-day challenges of parenting than mothers of children with similar demographic and social skills profiles (OR: 1.84; 95% CI: 1.28–2.65). Mothers with a child with autism were equally likely to have someone to whom they can talk about parenting on a day-to-day basis. Last, consistent with our findings in the univariate analysis, mothers with a child with autism dealt with disagreements about serious matters in the household just as other mothers did.

DISCUSSION
To our knowledge, this is the first population-based study of psychological functioning in mothers with a child with autism in the United States. Mothers of a child with autism had high levels of parenting stress and were more likely to report poor or fair mental and emotional health than mothers in the general population, even after adjustment for the child’s social skills and demographic background. This finding is consistent with the literature and suggests that other child or maternal factors, in addition to the child’s social skills and demographic background, may play a role in explaining the poorer mental health and increased stress of these mothers. However, we also found that mothers of a child with autism were indistinguishable from mothers in the general population with respect to having a close relationship, being angry with their child, or coping with parenting tasks. After adjustment for the child’s social skills and demographic background, mothers of a child with autism scored significantly better in all 3 areas. These results are in contrast to many negative findings in the

<table>
<thead>
<tr>
<th>TABLE 3 Multivariate Analyses of Maternal Functioning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Maternal health</td>
</tr>
<tr>
<td>Overall health (poor/fair)</td>
</tr>
<tr>
<td>Mental and emotional health (poor/fair)</td>
</tr>
<tr>
<td>Maternal Stress</td>
</tr>
<tr>
<td>Child is harder to care for than most children</td>
</tr>
<tr>
<td>Parent gives up her life to meet child’s needs</td>
</tr>
<tr>
<td>Parent–child relationship</td>
</tr>
<tr>
<td>Relationship is very close</td>
</tr>
<tr>
<td>Talk about important things with child</td>
</tr>
<tr>
<td>Angry at child</td>
</tr>
<tr>
<td>Child bothers mother a lot</td>
</tr>
<tr>
<td>Coping and support</td>
</tr>
<tr>
<td>Coping very well with parenting</td>
</tr>
<tr>
<td>Has someone to talk about parenting</td>
</tr>
<tr>
<td>Disagreement style</td>
</tr>
<tr>
<td>Keeps opinions to oneself</td>
</tr>
<tr>
<td>Discusses calmly</td>
</tr>
<tr>
<td>Argues heatedly</td>
</tr>
<tr>
<td>Hits or throws things</td>
</tr>
</tbody>
</table>

a OR for families with a child with autism compared with families with a child without autism, adjusted for demographic variables and child’s social skills. Demographic variables include age, gender, race, Hispanic ethnicity, poverty, education, and 2-parent family.
b P < .01.
c P < .001.
previous literature that compared mothers of children with autism with mothers of normally developing children. Furthermore, these findings raise questions regarding previous research that was based on referral-based samples that found that mothers of children with autism pursue psychological distancing as a coping strategy.  

Although mothers of a child with autism reported being bothered a lot by their child’s behavior with greater frequency than mothers in the general population in the univariate analysis, the association disappeared after controlling for demographic characteristics and the child’s social skills. This is consistent with previous research on a small number of preschool-aged children with autism that showed that maternal stress is influenced by the child’s behavior problems and suggests that the child’s poor social skills are an important reason for why mothers report the child as being bothersome.

Mothers with a child with autism reported that they were less likely to be able to talk to their child about things that really matter, even when compared with mothers of children with similar social skills. This result is not surprising given the communication impairments that are characteristic of autism. Consistent with this explanation, the level of the child’s social skills was positively associated with mothers’ being able to communicate with their child with autism (data not shown).

In the literature, the question of whether mothers of a child with autism had an “impaired capacity to believe that there are resources (parental support) available or that one has control of one’s parental skills” has been raised. In this study, we found that (1) mothers with a child with autism reported that they had someone with whom they can discuss the challenges of parenting on a day-to-day basis at comparable rates as the rest of the population and (2) mothers with a child with autism were more likely to report coping well with the challenges of parenting than mothers of children with similar demographics and social skills. Both of these findings contradict previous findings from less representative samples that mothers of children with autism have low parenting support and cope worse with parenting than other mothers.  

In this large, nationally representative sample of the United States, we also found no evidence that children with autism are less likely to live in 2-parent families. This result replicates previous research on a different nationally representative sample. Therefore, on a population basis, having a child with autism does not seem to increase the probability of the child living in a single-parent home. In addition, despite mothers’ having higher levels of stress and poor mental health, having a child with autism did not influence the way in which major disagreements were discussed in the household.

In summary, we found many strengths in the parent-child relationship, coping with parenting, and parenting support for mothers of a child with autism, in contrast to previous studies that used nonrepresentative samples. This suggests that some of the negative findings that were reported in the previous literature might stem from sampling bias.

Limitations
There are a few potential limitations in this study. First, all data were based on maternal report. Although parent report of autism is viewed as fairly reliable, it is possible that mothers with a child with autism report domains of psychological functioning differently than other mothers because mental health problems and a diagnosis of autism may influence recall bias. We did not have access to medical charts or diagnostic reports, and we were limited by the questions that were asked in the survey. In addition, some children in the sample may be undiagnosed or improperly diagnosed. Finally, information on personality characteristics that were shown to be related to our outcomes in another population was not available.

Implications
Although mothers of a child with autism experience high levels of stress and mental health problems, these mothers show remarkable strengths in coping, parent-child relationship, and psychological functioning. Given the challenges of parenting a child with autism, this suggests that families use compensatory strategies to maintain family stability in the context of poorer mental health and higher stress. A greater understanding of these potential strategies as well as the paths to resilient psychological functioning in the context of autism, stress, and poorer mental health problems is needed.

REFERENCES
7. Dunn ME, Burbine T, Bowers CA, Tantleff-Dunn S. Modera-
ARTICLE

Need for and Use of Family Leave Among Parents of Children With Special Health Care Needs

Paul J. Chung, MD, MS-a,b, Craig F. Garfield, MD, MAPP-c, Marc N. Elliott, PhD-d, Colleen Carey, BA-b, Carl Eriksson, MD, MPH-d, Mark A. Schuster, MD, PhD-a,b,e

aDepartment of Pediatrics, Mattel Children’s Hospital at UCLA, David Geffen School of Medicine at UCLA, Los Angeles, California; bRAND Health, RAND Corporation, Santa Monica, California; cDepartment of Pediatrics, Evanston Northwestern Healthcare, Feinberg School of Medicine at Northwestern University, Evanston, Illinois; dDepartment of Pediatrics, Children’s Hospital of Boston, Harvard Medical School, Boston, Massachusetts; eDepartment of Health Services, UCLA School of Public Health, Los Angeles, California

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

ABSTRACT

OBJECTIVE. Parents of children with special health care needs are especially vulnerable to work–family conflicts that family leave benefits might help resolve. We examined leave-taking among full-time–employed parents of children with special health care needs.

METHODS. We identified all children with special health care needs in 2 large inpatient/outpatient systems in Chicago, Illinois, and Los Angeles, California, and randomly selected 800 per site. From November 2003 to January 2004, we conducted telephone interviews with 1105 (87% of eligible and successfully contacted) parents. Among the sample’s 574 full-time–employed parents, we examined whether leave benefits predicted missing any work for child illness, missing >4 weeks for child illness, and ability to miss work whenever their child needed them.

RESULTS. Forty-eight percent of full-time–employed parents qualified for federal Family and Medical Leave Act benefits; 30% reported employer-provided leave benefits (not including sick leave/vacation). In the previous year, their children averaged 20 missed school/child care days, 12 doctor/emergency department visits, and 1.7 hospitalizations. Although 81% of parents missed work for child illness, 41% reported not always missing work when their child needed them, and 40% of leave-takers reported returning to work too soon. In multivariate regressions, parents who were eligible for Family and Medical Leave Act benefits and aware of their eligibility had 3.0 times greater odds of missing work for child illness than ineligible parents. Parents with >4 weeks of employer-provided leave benefits had 4.7 times greater odds of missing >4 weeks than parents without benefits. Parents with paid leave benefits had 2.8 times greater odds than other parents of missing work whenever their child needed them.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2337
doi:10.1542/peds.2006-2337

Key Words
chronic disease, family leave, parents

Abbreviations
CSHCN—children with special health care needs
FMLA—Family and Medical Leave Act
UCLA—Mattel Children’s Hospital at UCLA
CMH—Children’s Memorial Hospital
PedsQol—Pediatric Quality of Life Inventory
MHI-5—Mental Health Inventory 5
OR—odds ratio

Accepted for publication Nov 15, 2006
Address correspondence to Paul J. Chung, MD, David Geffen School of Medicine at UCLA, Department of Pediatrics, MOCC 12-325, 10833 LeConte Ave, Los Angeles, CA 90095.
E-mail: paulchung@mednet.ucla.edu

PEDIATRICS Volume 119, Number 5, May 2007 e1047
CONCLUSIONS. Full-time–employed parents of children with special health care needs experience severe work–family conflicts. Although most have leave benefits, many report unmet need for leave. Access to Family and Medical Leave Act benefits and employer-provided leave may greatly affect leave-taking.

FAMILY LEAVE PROGRAMS are intended to balance the demands of the workplace with the needs of families. Parents of chronically ill children or of children with special health care needs (CSHCN) are particularly vulnerable to difficulties with balancing work and family. CSHCN make up ~15% of US children and account for nearly half of US child health care expenditures. CSHCN average ~3 times as many medical encounters, hospitalizations, hospital days, and school absences as other children. During such medical encounters and hospitalizations, children and parents often need or want to be together, and when absent from school or child care as a result of illness, children need care or supervision. Therefore, employed parents of CSHCN may have a much greater need for time off than parents of healthy children.

Access to leave may be either provided by employers or created by federal or state legislation. Employer-provided leave is the major mechanism by which parents can take time off to care for ill children. Some employers grant leave to care for sick children, whereas some employees use their own sick leave or vacation, either surreptitiously (because their employers do not allow it) or in accordance with employer policies or state requirements. Parents with access to paid leave take more time off than other parents. However, parents of CSHCN tend to hold lower-paying jobs than parents of healthy children, and such jobs typically provide less access to paid leave and less ability to afford unpaid leave.

Federal legislation in 1993 established the Family and Medical Leave Act (FMLA), which provides eligible workers up to 12 weeks of unpaid leave per year to care for ill family members without risk for being fired for taking leave. Overall, it is believed that the FMLA has had only a modest impact on leave-taking. Only 47% of US employees are eligible, and many who are eligible cannot afford to take unpaid leave. In 2000, of the 3.5 million employees who said that they had recently needed family or medical leave but had not taken it, 78% cited inability to afford it as a reason. In July 2004, California became the first and only state to establish a paid family leave program, which aims to address some of the FMLA’s limitations; other states are exploring similar options.

In this study, we examined the availability and use of paid and unpaid leave among full-time–employed parents of CSHCN. We also identified factors that are associated with whether parents of CSHCN take leave and for how long. We hypothesized that access to the FMLA and employer-provided leave would be associated with greater amounts of work missed by these parents and greater self-reported ability to miss work whenever their child needed them.

METHODS

Sampling Frame

We sampled children who had received inpatient care, outpatient care, or both between October 1, 2002, and September 30, 2003, in 1 of 2 hospitals or their associated clinics: Mattel Children’s Hospital at UCLA (UCLA) or Children’s Memorial Hospital (CMH; Chicago, IL). Both sites are large, tertiary-care referral centers that serve economically and ethnically diverse populations.

The Maternal and Child Health Bureau defines CSHCN as children “who have or are at increased risk for a chronic physical, developmental, behavioral, or emotional condition and who also require health or related services of a type or amount beyond that required by children generally.” We adapted a previously validated approach to identify CSHCN using International Classification of Diseases, Ninth Revision billing codes. Researchers have developed a list of codes that are likely to be assigned to CSHCN. Not all patients who are identified in this manner, however, have enough health service needs to be considered CSHCN. Because there is a high correlation between illness severity and patient charges, we restricted the code list to disease categories with the highest average per-patient physician charges. These categories included bronchopulmonary dysplasia, cerebral palsy, chronic anemias, chronic enteritis/colitis, chronic renal failure, congenital heart diseases, cystic fibrosis, degenerative neurologic disorders, hydrocephalus, immunologic disorders, malignancies, organ transplant complications, and rheumatologic disorders.

We used this list to identify all children (younger than 18 years) in both hospitals’ billing databases who were (1) assigned 1 of these diagnoses at least once from October 1, 2002, to September 30, 2003, (2) listed as living, and (3) living in their respective states. We identified 1570 at UCLA and 3680 at CMH.

Stratified Random Sampling

Because our study focused on CSHCN with employed parents, we wanted to sample primarily from that group but without excluding children with poor or unemployed parents. The best proxy for employment/nonemployment status in the hospitals’ databases was Medicaid enrollment (among US adults aged 25–44 in 2004, 46% of Medicaid recipients versus 12% of others were not employed). We randomly sampled 800 children from each site, stratifying by their Medicaid status (100 Med-
icaid recipients and 700 Medicaid nonrecipients from each site) to achieve a smaller proportion of Medicaid recipients than in the CSHCN populations of both hospitals (33% Medicaid at UCLA, 31% at CMH). Inclusion of Medicaid recipients enabled inferences about how employment and leave decisions differed across the income spectrum.

**Participant Recruitment and Data Collection**

Trained interviewers telephoned the families, obtained informed consent by telephone, and conducted a 40-minute computer-assisted telephone interview in English or Spanish with 1 parent of each child. Our previous experience suggested that in 2-parent households, the mother or the parent who is employed fewer hours (often the same person) is more likely both to answer the telephone (because they are more likely to be home) and to consent to participate. To achieve better balance of employment and gender among household informants in participating 2-parent households (n = 915), interviewers asked whether both parents were available. When only 1 was available (n = 607), that parent was interviewed. When both were available (n = 308), interviewers asked whether 1 parent worked substantially more hours than the other. When 1 worked more hours (n = 207), interviewers attempted to enroll that parent. When both worked similar hours (n = 101), interviewers attempted to enroll the father. When the selected parent did not wish to participate (n = 108 of 308 2-parent households), the other parent was enrolled.

From November 2003 to January 2004 (ie, before the implementation of California’s paid family leave program), we completed interviews with 554 parents in California and 551 parents in Illinois. Excluding parents who were not located because of incorrect contact information (11% of all parents), who could not be reached despite repeated attempts (6%), or who were found to be otherwise ineligible (ie, the 3% who were younger than 18 years, who did not speak English or Spanish, who were too ill to participate, or whose child had died), completion rates were 85% at UCLA and 89% at CMH.

Parents who completed the interview received a $10 gift card. We received Health Insurance Portability and Accountability Act waivers to use hospital databases to obtain necessary private health information. The protocol was approved by all participating institutional review boards.

**Survey Content and Development**

Our survey covered parent job characteristics and benefits; parent need for and use of leave; relationships among child illness, parent employment, and parental leave; awareness and use of the FMLA; child quality of life; parent mental health; and family demographics. Some items were drawn from standardized instruments: (1) the Department of Labor 2000 Survey of Employees,13 (2) a short version of the Pediatric Quality of Life Inventory (PedsQL),20 and (3) the RAND Mental Health Inventory 5 (MHI-5).27 Other items were developed by the research team; reviewed by clinicians, attorneys, and social scientists who were familiar with CSHCN and/or labor issues; and pilot-tested on a convenience sample of parents. Items used Likert scales and close-ended response categories.

**Statistical Analysis**

**Construction of Weights**
Combining UCLA and CMH data, we compared responders (n = 1105) and eligible nonresponders (n = 165) across several variables (International Classification of Diseases, Ninth Revision category, child age, site, Medicaid status). Parents at UCLA were more likely to be nonresponders (15% vs 11%; P = .05); no other variable significantly predicted nonresponse. Because nonresponse differed only on the basis of site, we constructed poststratification weights by Medicaid status within each site and then gave each site equal weight by standardizing weights to mean 1 within sites. All descriptive and inferential statistics use this weight.

We performed 3 multivariate logistic regressions on the subsample of parents who were employed full time (n = 574). For some variables, categories were collapsed to achieve adequate cell sizes or to satisfy model assumptions that were informed by regression diagnostics. We evaluated which factors predicted, in the past 12 months, (1) whether parents missed any work because of their child’s illness (1: yes; 0: no), (2) whether parents missed >4 weeks of work because of their child’s illness (1: yes; 0: no), and (3) whether parents were always able to miss at least some work when their child needed them to (1: yes; 0: no).

**Missing Data**
No variable was missing for >3% of observations. To prevent the bias that can result from multivariate analyses that are restricted to complete cases,28 we used the multiple imputations by chained equations approach, a standard technique to impute missing data.29 The rate of missing data were sufficiently low that multiple imputations were unlikely to improve the efficiency of the imputations substantially, so we used only a single imputation from this method.30 In addition, the PedsQL is designed for children who are 2 years and older. To prevent dropping children who were younger than 2 years from the sample, we assigned them the mean PedsQL score and marked them with an indicator variable that estimated differences between those who were younger than 2 and those with the mean PedsQL score without biasing the estimate of PedsQL on outcomes.
For each of the 3 outcomes, we ran bivariate regressions for the following predictor variables, 1 at a time: 5 dichotomous indicators of access to and awareness of leave (availability of sick leave/vacation, availability of employer-provided leave other than sick leave/vacation, whether that leave was paid, FMLA eligibility, and awareness of FMLA eligibility), 8 parent demographic and health measures (respondent MHI-5 score, age, gender, race/ethnicity, education, marital status, household income, and partner’s employment status), and 6 measures of child health and demographics (child hospitalization frequency, PedsQL score, age, gender, Medicaid status, and site). To achieve both parsimonious and parallel parameterizations across the multivariate models, we set $P < .067 = .20/3$ in any of the bivariate regressions as the threshold for admission of a predictor variable into all 3 multivariate models. Using $P < .067$ preserved the $P < .20$ family-wise error rate that would have been standard for retention of a predictor variable if we had used just a single multivariate model.\(^{31}\) Doing so resulted in exclusion of marital status, partner’s employment status, child gender, and site from the final regression models. Interaction terms between site and leave-related predictor variables did not significantly improve the overall $\chi^2$ of the 3 regressions and were not retained.

## RESULTS

### Demographics of Responding Parents

Fifty-two percent of responding parents were female, and 79% were married or living as married (from this point, referred to simply as “married”; Table 1). Fifty-two percent were non-Hispanic white, 27% were Hispanic, 13% were non-Hispanic black, and 8% were “other.” Thirty-nine percent had annual household incomes less than $50,000, 26% had not attended college, and 19% of their CSHCN were Medicaid recipients. Among married respondents, 52% of partners were employed full time. The average MHI-5 score was 71 of 100. Thirty-six percent met usual criteria (MHI-5 $\leq 68$) for poor mental health, and 15% met usual criteria (MHI-5 $< 52$) for severe depressive symptoms.\(^{32}\)

### Child Illness Burden

On average in the past 12 months, children stayed home from school or child care 20 days because of illness, had 12 doctor or emergency department visits, had 1.7 hospitalizations, and (among the 41% who had any hospitalizations) spent 23 nights in the hospital (Table 1). The mean PedsQL score was 68 of 100.

### Leave-Taking Among Parents

Nineteen percent of parents took no time off in the previous 12 months to care for their ill child, 33% missed 1 week or less, 30% missed 1 to 4 weeks, and 18% missed $>4$ weeks (Table 2). During their longest leave, 60% received at least some pay from their employers.

### Ending Leave Prematurely

Fifty-five percent of responding parents said that, in the past 12 months, they were not always able to spend enough time with their child when he or she was ill. Of the parents who missed any work, 40% said that they returned to work sooner than was needed for their child’s health. The 3 most commonly cited reasons for returning to work were that the child was better (81%), the parents wanted to get back to work (66%), and the parents needed the pay (55%). Of the
TABLE 2  Employment Characteristics and Leave-Taking

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>% (N = 574)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>FMLA eligibility</td>
<td></td>
</tr>
<tr>
<td>Ineligible</td>
<td>51</td>
</tr>
<tr>
<td>Eligible, self-reported ineligibleb</td>
<td>18</td>
</tr>
<tr>
<td>Eligible, self-reported eligible</td>
<td>31</td>
</tr>
<tr>
<td>Access to employer-provided leave</td>
<td></td>
</tr>
<tr>
<td>Sick leave/vacation, d</td>
<td></td>
</tr>
<tr>
<td>0–5</td>
<td>20</td>
</tr>
<tr>
<td>6–20</td>
<td>46</td>
</tr>
<tr>
<td>&gt;20</td>
<td>34</td>
</tr>
<tr>
<td>Other leave, paid or unpaid, d</td>
<td></td>
</tr>
<tr>
<td>0–5</td>
<td>33</td>
</tr>
<tr>
<td>6–20</td>
<td>16</td>
</tr>
<tr>
<td>&gt;20</td>
<td>14</td>
</tr>
<tr>
<td>Parent does not know</td>
<td>37</td>
</tr>
<tr>
<td>Paid other leave, d</td>
<td>15</td>
</tr>
<tr>
<td>Work missed because of child’s illness (past year), d</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>1–5</td>
<td>33</td>
</tr>
<tr>
<td>6–20</td>
<td>30</td>
</tr>
<tr>
<td>&gt;20</td>
<td>18</td>
</tr>
<tr>
<td>During longest leave (past year)</td>
<td></td>
</tr>
<tr>
<td>Received any pay</td>
<td>60</td>
</tr>
<tr>
<td>Returned to work because needed income</td>
<td>55</td>
</tr>
<tr>
<td>Returned to work sooner than parent</td>
<td>40</td>
</tr>
<tr>
<td>On at least 1 occasion; did not miss work</td>
<td>41</td>
</tr>
<tr>
<td>even though parent believed needed to</td>
<td></td>
</tr>
<tr>
<td>miss because of child’s illness (past year)</td>
<td></td>
</tr>
<tr>
<td>Always able to spend as much time with ill</td>
<td>45</td>
</tr>
<tr>
<td>child as parent thought child needed (past year)</td>
<td></td>
</tr>
</tbody>
</table>

* Percentages are weighted to account for nonresponse and stratification.  
b Met criteria for eligibility (worked 1250 hours in the past 12 months, worked for the same employer for at least 12 months, and worked for an employer with at least 50 employees) but did not believe that they were eligible.

parents who returned to work even though their child was not better, 64% said that they returned because they needed the pay. Sixty-nine percent of parents said that they would have missed more work if they had received some or more pay during their time off.

Missed Opportunities for Leave

Forty-one percent of responding parents said that, in at least 1 instance in the past 12 months, they did not miss work even though they believed that they needed to because of their child’s illness. The 3 most commonly cited reasons for not missing work were that the parents could not afford to miss income (62%), they thought that they might lose their job or business (41%), and they thought that they might hurt their job advancement (37%). Seventy-three percent said that they would have missed work if they would have received some or more pay during that time off.

Access to Leave Predicts Leave-Taking

Access to the FMLA

Forty-nine percent of parents were FMLA eligible, satisfying all 3 criteria: worked at least 1250 hours in the past 12 months, worked for the same employer for at least 12 months, and worked for an employer with at least 50 employees (Table 2). Sixty-four percent of FMLA-eligible parents were aware that they were legally guaranteed time off. In bivariate regressions, parents who were eligible for the FMLA and aware of their eligibility were more likely than parents who were eligible but unaware (89% vs 72%; odds ratio [OR]: 3.1; P < .001) and ineligible (89% vs 79%; OR: 2.2; P = .005) to miss at least 1 day of work because of their child’s illness. This association persisted in multivariate regressions (Table 3). Parents who were eligible but unaware did not differ significantly from ineligible parents in bivariate or multivariate analyses.

Access to Sick Leave or Vacation

Eighty percent of parents had access to >1 week of employer-provided sick leave or vacation (34% had access to >4 weeks). Parents with access to sick leave/vacation did not differ significantly from parents without access in bivariate or multivariate regressions.

Access to Other Employer-Provided Leave (Paid or Unpaid)

Thirty percent of parents had access to other forms of employer-provided leave (paid or unpaid), such as family or personal time off. In bivariate regressions, parents with access to >4 weeks of other employer-provided leave were more likely than parents with no access to miss >4 weeks of work because of their child’s illness (32% vs 13%; OR: 3.1; P = .002). This association persisted in multivariate regressions.

Access to Other Employer-Provided Paid Leave

Of the parents who had access to other employer-provided leave, 50% had access to at least some pay during some or all of that leave. Contrary to our hypothesis, parents with access to paid leave were actually less likely in bivariate regressions to miss >4 weeks of work (8% vs 20%; OR: 0.34; P = .008) than parents without access. In multivariate regressions, however, this association was no longer significant (OR: 0.34; P = .11). Moreover, in multivariate regressions, parents with access to paid leave were more likely always to be able to miss at least some work when their child needed them to (OR: 2.8; P = .009) than parents without access.

Other Important Predictors of Leave-Taking

In bivariate and multivariate regressions, the number of hospitalizations was by far the best predictor of whether parents missed any work and whether they missed >4 weeks (Table 3). In bivariate regressions, women were
far more likely than men both to miss any work and to miss >4 weeks; the association between gender and missing >4 weeks persisted in multivariate regressions. Finally, higher MHI-5 (better parent mental health) and PedsQL (higher child quality of life) scores were strongly associated with less likelihood of missing any work or missing >4 weeks and greater likelihood of always being able to miss work when their child needed them to; in each case, 2 of the 3 associations persisted in multivariate regressions (Table 3).

**DISCUSSION**

Full-time–employed parents of CSHCN in our study experienced substantial conflict between work and family. Fewer than half reported always being able to spend enough time with their child when he or she was ill, and

---

**TABLE 3** Multivariate Regressions for Parent Missing Work Because of Child’s Illness and Parent’s Ability to Miss Work When Child Needed Them to

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Missed Any Work Because of Child’s Illness, OR (95% CI)</th>
<th>Missed &gt;4 wk of Work Because of Child’s Illness, OR (95% CI)</th>
<th>Always Able to Miss Work When Child Needed Them to, OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parent variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (unit = 5 y)</td>
<td>0.83 (0.65–1.1)</td>
<td>1.0 (0.97–1.6)</td>
<td>1.1 (0.91–1.3)</td>
</tr>
<tr>
<td>Female gender</td>
<td>1.2 (0.68–2.1)</td>
<td>5.2 (2.2–12)*</td>
<td>1.1 (0.68–1.8)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>0.70 (0.34–1.4)</td>
<td>0.64 (0.19–2.1)</td>
<td>1.5 (0.72–3.2)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1.1 (0.53–2.4)</td>
<td>1.3 (0.45–3.9)</td>
<td>0.39 (0.21–0.74)*</td>
</tr>
<tr>
<td>Other</td>
<td>0.29 (0.10–0.83)*</td>
<td>1.4 (0.50–4.0)</td>
<td>1.4 (0.67–2.9)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school graduate</td>
<td>0.40 (0.10–1.6)</td>
<td>1.6 (0.40–6.0)</td>
<td>3.1 (1.0–9.3)</td>
</tr>
<tr>
<td>High school graduate</td>
<td>1.0 (0.46–2.2)</td>
<td>0.50 (0.13–1.9)</td>
<td>1.1 (0.53–2.1)</td>
</tr>
<tr>
<td>Some college</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
</tr>
<tr>
<td>College graduate</td>
<td>0.70 (0.34–1.5)</td>
<td>0.26 (0.08–0.87)*</td>
<td>1.0 (0.53–1.9)</td>
</tr>
<tr>
<td>More than college graduate</td>
<td>0.88 (0.42–1.9)</td>
<td>0.26 (0.07–0.96)*</td>
<td>0.73 (0.38–1.4)</td>
</tr>
<tr>
<td>Annual household income, $</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20 000</td>
<td>0.63 (0.11–3.6)</td>
<td>1.4 (0.30–6.4)</td>
<td>1.9 (0.64–5.9)</td>
</tr>
<tr>
<td>20 000–49 999</td>
<td>0.84 (0.41–1.7)</td>
<td>0.75 (0.25–2.2)</td>
<td>0.97 (0.53–1.8)</td>
</tr>
<tr>
<td>50 000–99 999</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
</tr>
<tr>
<td>100 000–149 999</td>
<td>0.71 (0.36–1.4)</td>
<td>1.2 (0.39–3.8)</td>
<td>1.6 (0.86–3.1)</td>
</tr>
<tr>
<td>≥150 000</td>
<td>1.2 (0.57–2.7)</td>
<td>5.7 (1.4–24)*</td>
<td>1.6 (0.77–3.4)</td>
</tr>
<tr>
<td>MHI-5 score (unit = SD)</td>
<td>0.79 (0.56–1.1)</td>
<td>0.41 (0.29–0.58)*</td>
<td>2.1 (1.6–2.7)*</td>
</tr>
<tr>
<td><strong>Child variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrolled in Medicaid</td>
<td>2.9 (0.63–13)</td>
<td>6.7 (1.8–25)</td>
<td>0.73 (0.30–1.8)</td>
</tr>
<tr>
<td>Age (unit = 5 y)</td>
<td>0.75 (0.54–1.0)</td>
<td>0.69 (0.43–1.1)</td>
<td>1.1 (0.91–1.3)</td>
</tr>
<tr>
<td>Age &lt;2 y</td>
<td>0.24 (0.08–0.75)*</td>
<td>5.8 (1.6–21)*</td>
<td>3.1 (1.1–9.0)*</td>
</tr>
<tr>
<td>PedsQL score (unit = SD)</td>
<td>0.52 (0.35–0.76)*</td>
<td>1.2 (0.74–1.9)</td>
<td>1.6 (1.2–2.0)*</td>
</tr>
<tr>
<td>Hospitalizations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
</tr>
<tr>
<td>1</td>
<td>6.7 (2.8–16)*</td>
<td>9.7 (3.7–26)*</td>
<td>0.70 (0.39–1.3)</td>
</tr>
<tr>
<td>2–3</td>
<td>6.7 (1.9–24)*</td>
<td>15.5 (5.1–43)*</td>
<td>0.48 (0.22–1.0)</td>
</tr>
<tr>
<td>≥4</td>
<td>18 (2.2–150)*</td>
<td>70 (23–210)*</td>
<td>0.50 (0.22–1.1)</td>
</tr>
<tr>
<td><strong>Leave variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FMLA eligibility</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ineligible</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
</tr>
<tr>
<td>Eligible, self-reported ineligible</td>
<td>1.0 (0.53–2.0)</td>
<td>1.3 (0.44–3.7)</td>
<td>0.63 (0.35–1.1)</td>
</tr>
<tr>
<td>Eligible, self-reported eligible</td>
<td>3.0 (1.6–5.6)*</td>
<td>1.1 (0.46–2.4)</td>
<td>0.83 (0.49–1.4)</td>
</tr>
<tr>
<td>Access to sick leave/vacation, d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–5</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
</tr>
<tr>
<td>6–20</td>
<td>1.2 (0.53–2.6)</td>
<td>1.1 (0.39–3.2)</td>
<td>0.71 (0.36–1.4)</td>
</tr>
<tr>
<td>≥21</td>
<td>0.81 (0.34–1.9)</td>
<td>1.5 (0.45–5.2)</td>
<td>1.2 (0.56–2.4)</td>
</tr>
<tr>
<td>Access to other employer-provided leave, d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–5</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
<td>1.0 (reference)</td>
</tr>
<tr>
<td>6–20</td>
<td>0.97 (0.36–2.6)</td>
<td>0.45 (0.13–1.6)</td>
<td>0.64 (0.29–1.4)</td>
</tr>
<tr>
<td>≥21</td>
<td>0.72 (0.30–1.8)</td>
<td>4.7 (1.6–14)*</td>
<td>1.9 (0.91–4.0)</td>
</tr>
<tr>
<td>Parent does not know how many days</td>
<td>1.3 (0.61–2.6)</td>
<td>1.1 (0.47–2.7)</td>
<td>2.5 (1.4–4.5)*</td>
</tr>
<tr>
<td>Employer-provided leave is paid</td>
<td>0.97 (0.39–2.4)</td>
<td>0.34 (0.09–1.3)</td>
<td>2.8 (1.3–6.2)</td>
</tr>
</tbody>
</table>

CI indicates confidence interval.

*P < .05.
nearly half reported not missing work on at least 1 occasion even though their child needed them to. Although most parents had access to employer-provided leave, only 15% had access to paid leave outside sick leave/vacation. A majority of parents cited financial concerns as a major reason both for ending leave and for not taking leave even though their child needed them to. In multivariate regressions, whether parents missed any work was associated with their awareness of FMLA eligibility, whether they missed >4 weeks of work was associated with their access to employer-provided leave (paid or unpaid), and whether they always felt able to miss work was associated with their access to paid leave.

FMLA eligibility, awareness of FMLA eligibility, and access to employer-provided leave (both paid and unpaid) may have a substantial impact on the ability of parents of CSHCN to spend time with their ill children. In our sample, 49% of parents were eligible for the FMLA, and 64% of eligible employees were aware of the law. These numbers are very similar to national estimates of employees (47% eligibility and 58% awareness among eligible employees) and suggest that a large percentage of parents of CSHCN might benefit from both expansion of FMLA eligibility and increased education about the FMLA. Increasing access to employer-provided or state-provided leave, especially paid leave, may also provide benefits to these parents. California’s new Paid Family Leave Insurance program and similar programs that are being considered in other states might serve as models.

Moreover, it is possible that some parents of CSHCN in our sample had previously adjusted their employment in response to their child’s illness, thereby leading our results to reflect an already-optimized situation with respect to work–family conflict. Therefore, our results would underestimate the degree of conflict among parents of more newly diagnosed CSHCN. In addition, among the 52% of parents with full-time–employed partners, the work–family conflict that was reported by 1 parent may have underestimated the sum of conflicts that are experienced by both parents. In bivariate analyses of parents with full-time–employed partners, respondents whose partners missed more work than other respondents’ partners actually reported more conflict rather than less with respect to all 3 outcomes. The subsample, however, was too small to support multivariate analyses adequately. Future studies might examine the hypothesis that even though partners may share the burden of caring for ill children, partnership alone may provide only incomplete relief.

Although parents with access to paid leave were more likely than other parents to report being able to miss work, they were actually less likely to miss >4 weeks in bivariate analyses. Parents with access to paid leave might tend to hold positions of greater skill, responsibility, interest, or schedule flexibility, any of which could encourage or coerce employees to return more quickly. This hypothesis is supported by 2 additional findings (Table 3). First, more educated parents were less likely than other parents to miss >4 weeks. Second, parents whose children received Medicaid were more likely than other parents to miss >4 weeks. Therefore, our data may be describing an environment in which the parents with access to paid leave are currently predisposed to limiting their use of it. Whether this self-restraint would change in an environment of expanded access is unclear.

We found associations between higher parent MHI-5 scores and both more ability to miss work when their child needed them to and less likelihood of missing >4 weeks (Table 3). If one were to hypothesize that missing work should have a beneficial effect on parent mental health, then the latter finding might seem to contradict that hypothesis. Our observational cross-sectional study design, however, prevents us from examining whether more ability to miss work or missing less work increased parent mental health, whether better parent mental health reduced parents’ need to miss work, or whether some other factor (eg, severity of the child’s illness) influenced both. Longitudinal studies that examine change in MHI-5 scores over time would be useful. Regardless, MHI-5 scores among the parents of CSHCN in our study were substantially worse than in a previous nationally representative study of parents.31 In that study, only 17% of parents had MHI-5 scores <68, compared with 36% here.

We also found associations between higher PedsQL scores and both more ability to miss work and less likelihood of missing any work (Table 3). Again, our observational cross-sectional study design prevents us from examining the causal directionality of these associations. We can conclude, however, that the CSHCN in our study created a substantial illness burden for their parents. Our sample of children had an average quality-of-life score of 68. In previous studies, healthy, nonoverweight children had an average PedsQL score of ~85,26,34,35 children who were obese or had mild asthma had an average score of ~75,26,34 and children who had moderate to severe asthma or were receiving treatment for cancer had an average score of ~65,26,35

Because we selected children with higher expenditure diagnoses, excluding common but often relatively low-acuity chronic diseases such as attention-deficit/hyperactivity disorder or asthma, our children were somewhat sicker than children in nationally representative samples of CSHCN. CSHCN in the 2001 National Survey of Children with Special Health Care Needs averaged 7 missed school days per year (compared with 20 in our sample),26 and CSHCN in the 2000 Medical Expenditure Panel Survey averaged 10 doctor or emergency visits per year (compared with 12 in our sample).27 Nevertheless, the employment status of parents in our complete sample (80% of families had at least 1 parent
employed full time, 8% had at least 1 parent employed part time, and 12% had no parent employed) closely matched that of parents of CSHCN in the 2000 National Health Interview Survey (76%, 10%, and 14%, respectively), suggesting that our sample of parents may have been fairly representative of parents of CSHCN generally.

Full-time-employed parents of CSHCN experience work–family conflict that generates substantial stress on families and employers and seems to have the potential to affect child health and parent mental health. Future research should examine whether interventions that increase access to or awareness of employer-provided or state-provided leave, paid or unpaid, may help to alleviate this conflict and at what cost to employers and society. Research should also address the potential health effects of leave-taking on both children and parents and the potential role of health care providers in increasing awareness of leave policies among the most vulnerable families.

ACKNOWLEDGMENTS
We thank the Centers for Disease Control and Prevention (grant U48/DP000056), RAND Health, RAND Institute for Civil Justice, the California Endowment, and the National Institutes of Health (grant R21 HD052586) for financial support.

We thank Mattel Children’s Hospital at UCLA and Children’s Memorial Hospital in Chicago for administrative data services and cooperation with the study, the RAND Survey Research Group for interview data collection and management, the UCLA/RAND Work-Family Advisory Board for advice during questionnaire development, and the UCLA/RAND Center for Adolescent Health Promotion for staff and research assistance.

REFERENCES
16. Kingsbury K. Time off, with pay? Massachusetts leads a push across the U.S. to give paid family leave to every worker. Time. 2006;167:51
Sleep Duration and Overweight in Adolescents: Self-reported Sleep Hours Versus Time Diaries

Kristen L. Knutson, PhD, Diane S. Lauderdale, PhD

Department of Health Studies, University of Chicago, Chicago, Illinois

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

ABSTRACT

OBJECTIVE. Studies of children have found an inverse association between sleep duration and overweight on the basis of parental report of sleep duration. Studies in adolescents have been inconsistent but used varied measures of sleep. We used a nationally representative sample of adolescents that included 2 different measures of sleep duration (24-hour time diaries and self-reported usual sleep hours) to examine whether the association with overweight is sensitive to how sleep is measured in a single study population. We expected that the 2 measures of sleep would be strongly correlated and that the time-diary sleep would be more strongly associated with overweight risk because it is likely the more accurate measure of sleep.

METHODS. In 2001–2002, adolescents from the Child Development Supplement of the Panel Study of Income Dynamics completed 24-hour time diaries on a random weekday and weekend. Adolescents also self-reported average sleep duration. Both sleep measures were categorized into quartiles. Height and weight were measured, and a BMI z score for age and gender was calculated. Overweight was defined as above the 95th percentile.

RESULTS. The final sample included 767 male and 779 female subjects who were aged 10 to 19 years. Mean time-diary sleep was nearly 9 hours on weekdays and >10 hours on weekends. Mean self-reported sleep duration was 8 hours. Time-diary sleep and self-reported sleep were weakly correlated. Time-diary sleep was not significantly associated with overweight. Self-reported sleep was associated with overweight, but the association was not linear. When both sleep measures were included in the same model, their effects on overweight were independent.

CONCLUSIONS. The weak correlation between self-reported sleep and time-diary sleep and the independence of their associations with overweight raise questions about what different measures of sleep duration in adolescents represent.
Much research has been devoted to the examination of sleep among adolescents, and there are good reasons for this attention. First, the sleep of adolescents differs from that observed among prepubertal children and among adults. For example, the ability to fall asleep occurs later in the day, and adolescents typically obtain less sleep than younger children despite a similar sleep need of ~9 hours.1,2 Indeed, a recent national poll by the National Sleep Foundation reported that the average self-reported sleep duration of adolescents was 7.6 hours per night in 2005,3 over 1 hour less than their estimated sleep need. That adolescents obtain less sleep than they are thought to need is cause for concern because important negative outcomes have been associated with sleep loss, notably impaired memory and learning.4 Recent studies have demonstrated an association between sleep loss and impaired glucose metabolism and weight gain.5–8 The dramatic rise in the prevalence of overweight and obesity among US adolescents9 is an alarming development because obesity in adolescence is a strong predictor of obesity in adulthood10 as well as a risk factor for cardiovascular disease and all-cause mortality in adulthood.11 As such, the possibility that sleep loss may play a role in the increase of overweight adolescents has led some health researchers to examine this association.

Several studies have observed an association between short sleep duration and increased BMI or increased risk for being overweight among children. Studies in younger children aged 3 to 10 all used parental report of sleep duration, a measure whose validity has not been well established. Most of these studies were cross-sectional in design12–15; however, 2 were longitudinal.16,17 Nonetheless, the results are consistent, indicating that children whose parents report that their sleep falls in the shortest sleep category, which varied between studies, were at 1.5 to 5 times the risk for being overweight or obese compared with children in the longest sleep category, which also varied by study. Fewer studies have examined the association between sleep and weight status in adolescents (children aged 10 or older).18–22 Their findings are not as consistent as the studies of younger children. Whereas 2 smaller studies that used wrist actigraphy, an objective measure of sleep using a digital recording of movement, found large and significant associations between weight status and sleep, 2 nationally representative surveys that were based in the United States and Australia found an association only among male adolescents but assessed sleep by adolescent self-report. Therefore, it is unclear whether the inconsistent findings among adolescents are the result of differences between study populations or between sleep measurement methods.

How different modes of measuring adolescent sleep relate to each other has received little attention, and whether the mode of sleep assessment alters the apparent association with health outcomes such as weight status has not previously been investigated. There are 3 common ways of determining sleep duration in studies outside of a clinical setting: (1) a survey question about usual sleep duration, (2) a sleep log, and (3) wrist actigraphy. Sleep logs, routinely used in sleep medicine, ask patients to record the exact time when they turn out the lights to try to fall asleep and when they wake up in the morning. Wrist actigraphy has not been used extensively in larger population-based studies; it is relatively expensive. One previous study among high school students in Rhode Island did compare survey questions about usual sleep hours on weeknights and weekends with both a sleep log and actigraphy. It found moderate correlations for weeknights between self-reported usual sleep hours and sleep logs (Pearson’s r = 0.61) and between self-report and actigraphy (Pearson’s r = 0.53) but low correlations for weekends (Pearson’s r = 0.38 and 0.31, respectively).23

Our study used a large, nationally representative survey of adolescents and measured sleep in 2 different ways, both asking a survey question about usual sleep hours and also asking respondents to maintain 24-hour time diaries, which provide similar information to a sleep log. First, we examined whether time-diary sleep and self-reported sleep duration were similar. Second, we compared the associations between these 2 measures of sleep duration and overweight status.

### METHODS

#### Sample

These data come from the Child Development Supplement (CDS) of the Panel Study of Income Dynamics (PSID), which is a longitudinal study of a representative sample of US individuals and the families in which they reside.24 The PSID was begun in 1968 and has been collecting data on thousands of American families for over 35 years. In 1997, the PSID supplemented its core data collection with additional information on PSID parents and their 0- to 12-year-old children, called the CDS. Of the 2705 families selected for the CDS-I, 2394 (88%) participated, providing information on 3563 children. In 2002–2003, CDS recontacted families who were in CDS-I and remained active in the PSID panel as of 2001. CDS-II successfully reinterviewed 2017 (91%) families who provided data on 2908 children/adolescents who were aged 5 to 18 years. This analysis examines only adolescents (>10 years of age) and is restricted to data that were collected in 2002 because height and weight both were measured at this time.

#### Variables

Sleep duration is assessed using 2 different methods: time diaries and self-report. Time diaries provide detailed information concerning all activities during a 24-hour
period. Participants were asked to keep time diaries for 2 specified 24-hour days, beginning at midnight, on 1 randomly sampled weekday and 1 randomly sampled weekend day. The time diary asked in which activity the child was participating at midnight and at what time the activity ended. The next activity, which began at the time the first activity ended, was then recorded, and this continued until the 24-hour day was accounted for. This method provided a total amount of time spent sleeping across a 24-hour period. In this analysis, we used only nocturnal sleep time, and we created a weekly average by weighting by day of week (5/7 × weekday + 2/7 × weekend). Figure 1 presents the time diary that was used in this study. Adolescents were expected to complete the diaries themselves or to complete it together with a parent. Studies have indicated that calculating how much time has been devoted to a specific activity (eg, exercise, watching television) from a time diary is more accurate than asking respondents to estimate the total amount of time they spent on that activity in a specified 24-hour period. The second measure of sleep in this study is self-reported sleep duration, which comes from a single question. The adolescents were asked, “How many hours of sleep do you usually get a night?” The outcome in this study was overweight status. Because BMI changes with maturation, the definition of overweight for adolescents is not as standardized as for adults. Here we implemented a definition recommended by the Centers for Disease Control and Prevention. First, BMI was calculated using measured weight (in kilograms) over height (in meters squared) and was then transformed into a z score for age and gender using the “zanthro” program in Stata software (Stata Corp, College Station, TX). This function uses the LMS method (L is the power in the Box-Cox transformation, M is the median value, and S is the generalized coefficient of variation) and 2000 Centers for Disease Control and Prevention growth reference data. The BMI z scores were then transformed into percentiles, and overweight was defined as >95th percentile.

Because each day adds up to 24 hours, there may be inverse correlations between sleep and daytime activities (eg, adolescents who sleep more might watch less television). Therefore, we included as potential confounders the amount of time spent on daytime activities that might be associated with overweight status: television viewing, physical activity or exercise, and video game and computer use. These covariates were ascertained from the 24-hour time diaries and all are expressed in hours per day. Television viewing was coded as a separate activity category available from the diary. Six categories of sports or active leisure were coded from the time diaries. These categories included lessons (eg, swim, tennis, skating, gymnastics, martial arts, aerobics, music), team-based sports (eg, football, basketball; organized meets, games, or practice), individual sports (eg, tennis; golf; organized meets, games, or practice), exercise (playing sports or exercising that is not part of an organized event), out of doors (eg, hunting, fishing, boating, bicycling), and walking (walking, hiking, and jogging/running). These 6 categories were summed here to create a physical activity variable. Also, because music lessons were grouped with sports lessons, a second physical activity score was calculated by summing only 5 of the categories excluding lessons. Finally, the amount of time spent playing video games was added to the amount of time spent using a home computer.

Sociodemographic covariates included age, race/ethnicity of adolescent, family income, and head of household’s education level. Race/ethnicity was identified by the primary caregiver and was coded using dummy variables with the following categories: white, non-Hispanic; black, non-Hispanic; Hispanic; Asian; and other. Because of very low prevalence of overweight in the Asian group, we combined the Asian and other race/ethnicity groups. Family income was collected in 2001 about tax year 2000 and was log-transformed because of its highly skewed distribution. Zero and negative incomes were recoded to $1 before the log transformation. The education level of the head of household was collected in 2001.

<table>
<thead>
<tr>
<th>Time</th>
<th>What did your child do?</th>
<th>Time Began</th>
<th>Time Ended</th>
<th>If watching TV was that a videotape or TV program?</th>
<th>If TV, video, computer games, books: What was the name of the (program/video/game/book) child was (watching/playing/reading)?</th>
<th>Where was child?</th>
<th>Who was doing the activity with child?</th>
<th>Who (else) was in the same location (see column F) but not directly involved in the activity</th>
<th>What else was child doing at the same time?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midnight</td>
<td>12:00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 1**
Example of the time diary used in the 2002 CDS of the PSID. Adolescents were expected to complete the diary themselves or together with a parent or caregiver.
and represented the total number of grades completed by the head of household, ranging from 0 to 17.

**Statistical Methods**

We excluded cases that were missing values for height or weight, weekday or weekend time-diary sleep, or self-reported sleep duration. We calculated the quartiles for both the time-diary sleep weighted by day of week and the self-reported sleep duration. A Pearson’s correlation coefficient was calculated to determine the bivariate association between time-diary and self-reported sleep as continuous variables, and a Spearman’s rank correlation coefficient was calculated to determine the bivariate association between the quartiles of time-diary and self-reported sleep. To understand further the association between sleep measures, we modeled self-reported sleep as a function of time-diary sleep using linear regression and adjusting for covariates.

Logistic regression models were then used to predict overweight status from the quartile sleep variables, which were entered as indicator variables with the highest quartile as the reference group. The first model included time-diary sleep, the second included self-reported sleep duration, and the third model included both measures of sleep duration. All models were adjusted for the covariates (television viewing, physical activity, video game/computer use, age, gender, race, income, and head of household education). Wald tests on the time-diary and self-reported sleep variables were performed for each regression analysis. Interaction terms between gender and each sleep quartile were added to test whether associations differed by gender. Models also were run separately by gender to examine differences in the odds ratios. Regression models included sample weights. Because more than 1 adolescent was sampled from some families, regression analyses were adjusted for clustering using family identification numbers. We performed all statistical analyses by using Stata 9.0 statistical software (Stata Corp).

**RESULTS**

The final sample, which excluded adolescents with missing data for BMI ($n = 32$), time-diary sleep ($n = 144$), or self-reported sleep duration ($n = 29$), included 767 male and 779 female subjects who were aged 10 to 19 years. Average age was 14.3 years (Table 1). Mean time-diary sleep was 8.8 hours on weekdays and 10.3 hours on weekends; however, mean self-reported sleep duration was only 8.0 hours. Figure 2 presents the distributions of self-reported and time-diary sleep. The modal response was 8 hours for self-reported sleep and an average of 9 hours for time-diary sleep. The Pearson’s correlation coefficient between the 2 measures was 0.27 ($P < .001$), and the Spearman’s rank correlation coefficient between the 2 quartile variables was 0.33 ($P < .001$). The $\beta$ coefficient for time-diary sleep from the adjusted regression analyses that predicted self-reported sleep was 0.20 ($P = .002$), indicating that every additional hour of time-diary sleep was associated with 12 minutes of additional self-reported sleep.

In all of the logistic regression models that predicted overweight, the referent sleep category is the longest sleep quartile. For the entire sample (Table 2), whether time-diary sleep (model 1) or self-reported sleep (model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>All ($N = 1546$)</th>
<th>Male Subjects ($n = 767$)</th>
<th>Female Subjects ($n = 779$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>14.3 (2.6)</td>
<td>14.2 (2.6)</td>
<td>14.4 (2.5)</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>23.0 (6.2)</td>
<td>22.8 (6.0)</td>
<td>23.2 (6.5)</td>
</tr>
<tr>
<td>BMI z score</td>
<td>0.60 (1.15)</td>
<td>0.61 (1.19)</td>
<td>0.59 (1.10)</td>
</tr>
<tr>
<td>Time-diary sleep, weekday, h</td>
<td>8.8 (1.6)</td>
<td>8.9 (1.6)</td>
<td>8.7 (1.7)</td>
</tr>
<tr>
<td>Time-diary sleep, weekend, h</td>
<td>10.3 (2.0)</td>
<td>10.3 (2.1)</td>
<td>10.4 (2.0)</td>
</tr>
<tr>
<td>Time-diary sleep, weighted by day of week, h</td>
<td>9.2 (1.4)</td>
<td>9.3 (1.4)</td>
<td>9.2 (1.4)</td>
</tr>
<tr>
<td>Self-reported sleep duration, h</td>
<td>8.0 (1.7)</td>
<td>8.1 (1.8)</td>
<td>7.9 (1.6)</td>
</tr>
<tr>
<td>Difference between time-diary and self-reported sleep, h</td>
<td>1.3 (1.9)</td>
<td>1.2 (2.0)</td>
<td>1.3 (1.9)</td>
</tr>
<tr>
<td>Television viewing, h</td>
<td>2.2 (1.7)</td>
<td>2.3 (1.7)</td>
<td>2.2 (1.6)</td>
</tr>
<tr>
<td>Physical activity, h</td>
<td>0.6 (0.9)</td>
<td>0.8 (1.0)</td>
<td>0.4 (0.8)</td>
</tr>
<tr>
<td>Video game/computer use, h</td>
<td>0.9 (1.4)</td>
<td>1.2 (1.6)</td>
<td>0.5 (1.0)</td>
</tr>
<tr>
<td>Income, $1000s</td>
<td>68.5 (93.7)</td>
<td>65.2 (75.2)</td>
<td>71.7 (108.9)</td>
</tr>
<tr>
<td>Head of household education, y</td>
<td>12.9 (2.7)</td>
<td>12.9 (2.6)</td>
<td>12.8 (2.8)</td>
</tr>
<tr>
<td>Overweight, %</td>
<td>21</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td>Race, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>46</td>
<td>45</td>
<td>47</td>
</tr>
<tr>
<td>Black</td>
<td>42</td>
<td>44</td>
<td>40</td>
</tr>
<tr>
<td>Hispanic</td>
<td>7</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Asian</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Data are means (SD) except where noted.
2) is the measure of sleep, there is a nonlinear pattern of odds for overweight by sleep quartiles. The shortest and longest quartiles have similar odds for overweight, whereas there is a higher and similar odds for the 2 middle quartiles. The elevated odds for the 2 middle quartiles is statistically significant for self-reported sleep (P = .05 and .02, respectively), whereas the elevated odds do not achieve statistical significance for time-diary sleep (P = .10 and .07, respectively). The Wald tests for the equivalence of the coefficients for the sets of sleep variables similarly are significant for self-reported sleep (P = .002) and not for time-diary sleep (P = .11). Inclusion of both time-diary and self-reported sleep in the same model (model 3) has almost no effect on the odds ratios or significance tests for either measure. The independence of their associations suggests that the association of each with overweight status is not related to their correlated information about underlying sleep habit.

Television viewing is significantly associated with risk for overweight. Every hour of television is associated with ~20% increased risk. Neither physical activity nor video game/computer use was significantly associated with risk for being overweight. Substituting the 5-category physical activity variable that excluded lessons did not alter any of the results.

None of the interaction terms between gender and the sleep quartiles was significant in any of the models (data not shown), indicating that the associations between the sleep measures and overweight status were not significantly different for male and female adolescents. When stratified by gender, however, the associations between time-diary sleep and overweight were stronger for female adolescents, whereas the associations between self-reported sleep and overweight were stronger in male adolescents. For female adolescents, the risk for overweight in the shortest self-reported sleep quartile was significantly reduced relative to the longest sleep quartile.

**DISCUSSION**

In this sample, the correlation between time-diary sleep and self-reported sleep was weak. Average time-diary sleep was >1 hour longer than self-reported sleep duration. The time-diary method used 2 random days that may not be representative of habitual behavior and that might account in part for the low correlation between the measures. However, that would not explain the large difference between the means of the 2 types of sleep data, unless participants systematically spent more time in bed because they were keeping the time diary. Furthermore, time diaries have in general been found to be more reliable than a summary estimate of time spent in a specific activity. Because we found that each additional hour of time-diary sleep was associated on average with only one fifth of an hour of additional self-reported sleep, it is unclear how to interpret the self-reported sleep duration variable. That the 2 sleep measures did not confound each other in the regression that predicted overweight further suggests that the 2 variables represent different aspects of sleep or that self-reported sleep duration also reflects factors other than sleep duration, such as sleep quality, psychosocial factors, or perceptions of socially desirable sleep habit.

Two previous nationally representative studies among adolescents that used self-reported sleep duration or time spent in bed found a significant association between sleep duration and BMI or risk for overweight in male but not female adolescents. Two other studies among adolescents used a more objective measure of sleep, wrist activity monitoring or actigraphy, in samples that were either clinic based or nonrepresentative. One study recruited overweight adolescents from a hospital-based weight management clinic and compared them with healthy control subjects from another study. They observed that overweight adolescents had a shorter nocturnal sleep period than the lean control subjects (7.8 vs 8.5 hours, based on actigraphy). The second study that used actigraphy measured total sleep time during a single 24-hour period in a sample of 383 adolescents, and an astonishingly strong association was observed between total sleep time and risk for obesity: an odds ratio of 0.20 predicting risk for obesity for each additional hour of sleep. Associations this strong have not been observed in other studies. To date, studies that have examined the association between overweight status and sleep have been inconsistent, but the studies have differed in important ways: how sleep is measured and whether the study population is clinic, volunteer, or community based.

In our study, the pattern and the magnitude of the effects generally were similar between the time-diary and self-report quartiles, with elevated risk for overweight for the 2 middle quartiles of sleep duration. The associations reached statistical significance for self-reported sleep only. The nonlinear pattern that we observed for both time-diary and self-reported sleep quartiles was not reported previously; neither has the low risk for overweight for the shortest sleepers. Insofar as
we did find significantly elevated risk for overweight for the middle 2 quartiles compared with the longest quartile, our findings generally are consistent with other studies among children and adolescents that examined the association between self- or parent-reports of sleep duration and BMI or overweight/obesity.12–17 Gender differences that are consistent with previous studies that used self-reported sleep hours were observed in that the associations between self-reported sleep quartiles and odds for overweight were stronger for male than for female adolescents; however, the interaction terms with gender were not significant.

The observation that self-reported sleep duration is more strongly associated with BMI than time-diary sleep is puzzling, as is their independence when both are included in the same model. It is possible that the 24-hour time diaries, which were obtained on 2 random days (1 weekday and 1 weekend), are not representative of typical sleep behavior, which the self-reported measure may better represent. However, the strongest association that was reported previously between sleep duration and obesity was observed in the study that used a single 24-hour period of sleep recording.19 One possibility is that factors other than sleep alone, such as psychological or behavioral characteristics, contribute to how an adolescent answers a question about how much sleep he or she usually gets and that some of these other factor are themselves associated with BMI. That self-reported sleep duration remained a significant predictor of overweight status after adjustment for time-diary sleep supports this possibility. Our results raise questions about what self-reported sleep duration for adolescents represents; additional research to assess the validity and the determinants of adolescent self-reported sleep duration is needed.

There are limitations to this analysis. Time that is noted in a time diary as spent in bed sleeping is not necessarily equivalent to actual sleep duration because it may not take into account time to fall asleep or wakeful periods during the night. Also, the 24-hour time diaries cover a period from midnight to midnight, which may differ from the amount of night time sleep recorded in a 24-hour period that includes a single night rather than parts of 2 sequential nights. Finally, this was a cross-sectional analysis, so no direction of causality between sleep and overweight may be inferred.

**CONCLUSIONS**

The prevalence of childhood and adolescent obesity is increasing at an alarming rate in the United States and worldwide, and sleep has been identified as a possible contributory factor. As previously reviewed, many epidemiologic studies among children from around the world support the hypothesis that sleep is associated with weight gain, and laboratory studies in adults have revealed a potential mechanism for this association.14–17

**TABLE 2**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1, Time-Diary Sleep, OR (95% CI)</th>
<th>Model 2, Self-reported Sleep, OR (95% CI)</th>
<th>Model 3, Both Sleep Variables, OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-diary sleep, h</td>
<td>3.56–8.38</td>
<td>1.02 (0.57–1.84)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>8.39–9.24</td>
<td>1.56 (0.92–2.64)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>9.25–10.07</td>
<td>1.63 (0.96–2.78)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>10.08–16.17</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Equality of time-diary groups, P</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Self-reported sleep, h</td>
<td>0.5–7.0</td>
<td>—</td>
<td>0.88 (0.45–1.69)</td>
</tr>
<tr>
<td></td>
<td>7.1–8.0</td>
<td>—</td>
<td>1.85 (1.01–3.38)</td>
</tr>
<tr>
<td></td>
<td>8.1–9.0</td>
<td>—</td>
<td>1.93 (1.10–3.37)</td>
</tr>
<tr>
<td></td>
<td>9.2–19.0</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Equality of self-reported sleep groups, P</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Television viewing, h/d</td>
<td>1.21 (1.09–1.34)</td>
<td>1.20 (1.08–1.32)</td>
<td>1.20 (1.08–1.33)</td>
</tr>
<tr>
<td>Physical activity, h/d</td>
<td>1.07 (0.85–1.35)</td>
<td>1.04 (0.83–1.30)</td>
<td>1.05 (0.83–1.31)</td>
</tr>
<tr>
<td>Video game/computer use, h/d</td>
<td>1.01 (0.85–1.18)</td>
<td>1.01 (0.86–1.18)</td>
<td>1.01 (0.86–1.19)</td>
</tr>
<tr>
<td>Age</td>
<td>0.96 (0.90–1.03)</td>
<td>0.97 (0.90–1.04)</td>
<td>0.96 (0.90–1.03)</td>
</tr>
<tr>
<td>Age²</td>
<td>1.00 (1.00–1.00)</td>
<td>1.00 (1.00–1.00)</td>
<td>1.00 (1.00–1.00)</td>
</tr>
<tr>
<td>Gender (0 = male; 1 = female)</td>
<td>0.65 (0.44–0.96)</td>
<td>0.66 (0.45–0.97)</td>
<td>0.67 (0.45–0.98)</td>
</tr>
<tr>
<td>ln(income)</td>
<td>1.04 (0.88–1.22)</td>
<td>1.04 (0.89–1.23)</td>
<td>1.04 (0.88–1.22)</td>
</tr>
<tr>
<td>Head of household education</td>
<td>0.95 (0.88–1.02)</td>
<td>0.96 (0.89–1.04)</td>
<td>0.96 (0.89–1.03)</td>
</tr>
<tr>
<td>Race</td>
<td>1.83 (1.15–2.91)</td>
<td>1.90 (1.19–3.02)</td>
<td>1.91 (1.19–3.06)</td>
</tr>
<tr>
<td>Black</td>
<td>1.83 (1.15–2.91)</td>
<td>1.90 (1.19–3.02)</td>
<td>1.91 (1.19–3.06)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.88 (0.43–1.81)</td>
<td>0.90 (0.44–1.85)</td>
<td>0.89 (0.43–1.85)</td>
</tr>
<tr>
<td>Asian/other</td>
<td>0.55 (0.22–1.42)</td>
<td>0.57 (0.21–1.50)</td>
<td>0.59 (0.23–1.53)</td>
</tr>
<tr>
<td>White</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
</tr>
</tbody>
</table>

a P < .05.
However, the validity of self-reported sleep duration, which is the primary measure of sleep used in epidemiologic studies, needs to be investigated for adolescents to determine whether results that were observed in these studies represent true associations. Clinicians who are interested in assessing habitual sleep behavior among adolescents may want to consider using a sleep diary in addition to a single self-reported measure of sleep duration.

ACKNOWLEDGMENT
Research for this study was supported by National Institutes of Health grant 1 R01 HL082907-01.

REFERENCES
Duration of Poverty and Child Health in the Quebec Longitudinal Study of Child Development: Longitudinal Analysis of a Birth Cohort

Louise Séguin, MD, MPH, Béatrice Nikiéma, MD, MSc, Lise Gauvin, PhD, Maria-Victoria Zunzunegui, PhD, Qian Xu, MD, MSc

Department of Social and Preventive Medicine and Interdisciplinary Health Research Group, University of Montreal, Montreal, Quebec, Canada

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

ABSTRACT

OBJECTIVES. The objective of this study was to examine the relationship between duration of poverty and the health of preschool children in the Quebec Longitudinal Study of Child Development birth cohort.

METHODS. Data from the Quebec Longitudinal Study of Child Development for 1950 children who were followed annually up to age 3½ years were analyzed. Poverty was defined as having an income below the low-income cutoff from Statistics Canada. Five health indicators were examined: asthma attacks, infections, growth delay, a cumulative health-problems index, and maternal perception of the child’s health. The association between duration of poverty and child health was explored with logistic regression modeling controlling for child and mother characteristics, including the mother’s level of education, social support, and physical violence.

RESULTS. In this birth cohort, 13.7% (268) 3½-year-old children from the Quebec Longitudinal Study of Child Development experienced intermittent poverty since birth (1–2 episodes), and another 14.4% (280) experienced chronic poverty (3–4 episodes). Children from families with chronic poverty had more frequent asthma attacks and had a higher cumulative health-problems index score, whereas children with intermittent poverty were more often perceived to be in less than very good health by their mothers. These associations remained statistically significant when controlling for child and mother characteristics. No association was observed between duration of poverty and infections or growth delay.

CONCLUSIONS. Chronic poverty affects a large number of children and has negative consequences for preschool children’s health, although universal health care is available. The effects of chronic poverty may vary according to different health indicators and the age of the child.
The relationship between socioeconomic status (SES) and child health has been observed in industrialized countries. These associations also are present when poverty is defined as a lack of material resources or low income. Poverty is not a static condition, and there are many entries in and exits from poverty in any 1 year. However, poverty dynamics are not always taken into account when the associations with children health are studied, because most studies in this area rely on cross-sectional measures of poverty. Short duration of poverty does not entail the same consequences for a child’s health as does longer duration of poverty. Chronic poverty seems to be especially harmful for young children because it is associated with a higher risk for mortality and for morbidity. Moreover, poverty during early childhood not only affects child health but also jeopardizes future adult health. Numerous adult chronic health problems, such as cardiovascular diseases, originate during pregnancy and the first years of life when the child’s family is poor. However, few longitudinal studies have examined the health of preschool children in relation to their family’s financial difficulties.

Population studies about child health most often analyze chronic health problems, asthma, infections, and growth delay. Most of these studies are cross-sectional, analyze only 1 or 2 health indicators, and include heterogeneous samples of children who are younger than 18 years. Few studies have controlled for both maternal and child characteristics that could attenuate the relationship between chronic poverty and child health. Last, most of these studies pertain to children who live in the United States. Universal access to health care and various social programs that are specific to Quebec, such as the low-cost child care system, could buffer the effects of poverty on health among poor children. In this article, we examine the relationships between duration of exposure to poverty since birth and the health of 3½-year-old children focusing on 5 health indicators (asthma attacks, infections, growth delay, a cumulative health-problems index [CHPI], and maternal perception of the child’s health) while controlling for the child’s and the mother’s health and social characteristics.

METHODS

Data came from the first 4 cycles of the Quebec Longitudinal Study of Child Development (QLSCD), which was coordinated by the Direction Santé Québec of the Institut de la statistique du Québec. A birth cohort of 2120 children was followed up annually since 1998. The sample was representative of singleton live births registered in the Quebec live births registry in 1997–1998 with the exception of those on the Cree and Inuit territories, on Indian reservations, or in the Northern region of Quebec (2.1% of live births). Infants who were born before 24 or after 42 gestational weeks (0.1%) and those with unknown gestational age (1.3%) were excluded. At 41 months, the participation rate was 92% (n = 1950).

Data on health conditions at birth were taken from hospital files. After informed consent was given, structured interviews were conducted at home by trained interviewers at 5, 17, 29, and 41 months with the person most knowledgeable about the child (the mother in 98% of cases). At 29 and 41 months, the interviewers measured the child’s height with a standardized procedure.

Five health indicators were used to examine the health of the child at 41 months: reports by the mother of the occurrence in the previous 12 months of an asthma attack, maternal report of the occurrence of an infection (respiratory, otitis, gastroenteritis, other) in the previous 3 months, perception of the child’s health by the mother, growth delay, and the CHPI. Growth delay was defined as a z score of the child’s height under the 10th percentile of the Centers for Disease Control and Prevention’s growth curves. The CHPI identifies children who accumulate 2 or more of the following health problems: asthma attacks, infections, and growth delay.

Poverty was defined in terms of insufficient household income, that is, having an income during the past 12 months below the Canadian low-income cutoff (LICO) as estimated by Statistics Canada. A family is said to be under the LICO when they attribute 20% more than the average Canadian family to food, shelter, and clothing. There are different cutoffs according to the number of people in a household and whether the household is located in a rural area or a small or large urban area. Duration of poverty was categorized as never poor, intermittently poor (1–2 episodes of being under the LICO out of the 4 follow-up periods), or chronically poor (3–4 episodes).

The possible confounding variables included child characteristics (gender, age, birth order, duration of breastfeeding, and birth conditions [preterm birth, small for gestational age, and congenital malformation]), mother characteristics (age at birth of the child, education, immigration status, height, and social support), and child environment (smoking in the house, family type, child care, and exposure to physical violence). Social support was assessed with a validated shortened version of the Social Provisions Scale, which originated from the National Longitudinal Study of Children and Youth. In particular, factor analysis of the National Longitudinal Study of Children and Youth data showed that 3 items could adequately ascertain social support: extent of mother’s access to people who make her feel secure and happy, extent of mother’s access to people in whom she can confide, and extent of mother’s access to people who can provide material aid when necessary. Summation of values on the 3 items was standardized to create a score that ranged from 0 to 10, with higher values indicating greater support. Physical violence was
established with 1 question that was addressed to the mother at 41 months, which inquired about whether she had been beaten, pushed, or hurt since the birth of the child. This question was adapted from the validated Abuse Assessment Screen.\textsuperscript{33}

In analysis, first we examined univariate associations between duration of poverty and each health indicator. Second, the possible confounding role of each risk factor was examined for each health outcome by fitting a logistic model. A risk factor was considered a confounder in the association of poverty with child’s health when the value of the coefficient of poverty changed by at least 10% after inclusion of the potential confounder.\textsuperscript{34} Third, a final model that included duration of poverty and all confounders was built. We also tested the potential interacting and mediating roles of social support in the poverty–health associations. All of the analyses were performed with SPSS 11.00 (SPSS, Chicago, IL). Missing cases were excluded listwise. This study was approved by the University of Montreal Faculty of Medicine Human Research Ethics Committee.

RESULTS

Results show that 13.7\% (268) of 3½-year-old Quebec children in the QLSCD study had experienced 1 or 2 episodes of poverty since their birth, and 14.4\% (280) had experienced chronic poverty with 3 or 4 episodes of poverty since their birth, so 28.1\% (548) of them had experienced chronic poverty with 3 or 4 episodes of poverty during their first years of life. More than one quarter (28.1\%) of families lived for at least 1 episode of poverty during their first years of life. That many children live in families who experienced at least 1 year with insufficient income between the time the child was born until the age of 3½ years. Table 1 shows the number of people in the household and in the residency area.

In the multivariate logistic regression analyses (Tables 4 and 5), poverty increased the odds of asthma attacks, greater CHPI scores, and perception of poor health by the mother. After adjustment for relevant confounding variables, children from chronically poor families have a greater probability of enduring asthma attacks than children from nonpoor families (odds ratio [OR]: 2.36; 95\% confidence interval [CI]: 1.31–4.25), and they are likely to present with >1 health problem according to the CHPI (OR: 1.87; 95\% CI: 1.13–3.09). Children with intermittent poverty have a higher probability of being perceived as being in less than very good health by their mothers (OR: 1.90; 95\% CI: 1.26–2.84), but the association becomes nonsignificant among children who are chronically poor (OR: 1.3; 95\% CI: 0.8–2.1).

Nonsignificant relationships between chronic poverty and infections (OR: 1.1; 95\% CI: 0.8–1.6) or growth delay (OR: 1.3; 95\% CI: 0.7–2.2) are observed, although estimated ORs are in the expected direction. We did not find a statistically significant interaction between social support and chronic poverty for any of the health outcomes considered.

In the asthma model, gender and birth order of the child remained significant predictors. In the CHPI model, the child’s gender and age, being born small for gestational age, the mother’s height, and the use of a child care center contributed uniquely to a poorer health outcome. In the analysis of maternal perception of the child’s health, low social support increased the odds of being perceived as being in less than very good health by the mother. Use of a child care center, being a non-European immigrant, and receiving low social support are linked with infections. Last, being born small for gestational age and short maternal height are predictors of growth delay.

DISCUSSION

The objective of this study was to examine the relationship between duration of exposure to poverty since birth and preschool children’s health among children in the QLSCD using 5 different health indicators. Results show that many children live in families who experienced at least 1 episode of poverty during their first years of life. More than one quarter (28.1\%) of families lived for at least 1 year with insufficient income between the time the child was born until the age of 3½ years. Moreover, for 14.4\% of them, this situation was repeated 3 to 4 times. This level of poverty is similar to what is observed in the rest of Canada and in other industrialized countries such as the United States and Great Britain, where children are most often afflicted by poverty among all age groups in the population.\textsuperscript{4,6,35,36} However, some European countries, using diverse fiscal measures, are successful in reducing the number of children who are exposed to poverty. According to a United Nations Children’s Fund report based on Organisation for Economic Co-operation and Development data,\textsuperscript{37} af-

\begin{table}[h!]
\centering
\begin{tabular}{ll}
\hline
No. of Episodes of Poverty & \% (n) \\
\hline
0 & 66.5 (1297) \\
1–2 & 13.7 (268) \\
3–4 & 14.4 (280) \\
Missing & 5.4 (105) \\
\hline
\end{tabular}
\caption{Distribution of 3½-Year-Old Children According to Duration of Exposure to Poverty}
\end{table}

Poverty is defined according to Statistics Canada low-income thresholds, which take into account the number of people in the household and in the residency area.
ter taxes and social benefits, 14.9% of Canadian children were living in a poor family in 2000, 21.9% of children were poor in the United States, but no more than 2.8% to 4.2% of children were living in poverty in Scandinavian countries.

Chronic poverty has negative impacts on the health of preschool children who have more frequent asthma attacks and a greater number of health problems than children from nonpoor families. These results confirm those that link health with chronic poverty. Some authors hypothesized that the more frequent asthma attacks among poor children would be the result of limited access to health care, yet here we report a similar finding for children who live in a society with universal access to health care and a low-cost child care system. Similar results were observed in other parts of Canada as well as in other countries where health care is financially accessible. In this cohort, we observed that poor mothers consulted health professionals for their children as often as did nonpoor mothers, yet it is possible that mothers in poor families have more difficulty complying with the daily treatment for their child’s asthma. Unfavorable housing conditions, living in a more polluted environment, and more frequent smoking by their parents can play a role in eliciting more frequent asthma attacks among chronically poor children. Moreover, stress that results from living in conditions of poverty might play a major role in the links between low SES and asthma in children.

Chronic poverty is also associated with a higher score on the CHPI, which means that poor children are more likely to have >1 of the health problems included in the index. Similar results were observed in other studies that used a similar index. This higher number of health problems among poor children might reflect a higher general vulnerability that could result from chronic stress. Stress that results from these difficult living conditions and the biological stress process that ensues can affect the body’s capacity to prevent onset of disease processes.

In this analysis, maternal perception of children’s health is worse only for children who experienced transient episodes of poverty. In an earlier analysis, we demonstrated that maternal perception of child’s health is a valid indicator of health at 17 months because it was associated with the presence of health problems in the child after controlling for mother’s demographic and socioeconomic characteristics. That maternal perception of child health is not worse for children who are chronically poor compared with children who were never poor might result from a lack of statistical power. It also could be attributable to the selected attrition of mothers with a low level of education, mothers who head single-parent families, and non-European immigrant mothers who also are chronically poor mothers.

In this analysis, infections were not associated with chronic poverty at this age. In a previous analysis of these children when they were aged 2½ years, a clear link between chronic poverty and infections was reported. At 3½ years, more children regularly attend a child care center, a widely known source of infections. However, children from families who did not experience poverty are more often cared for in a child care center than poor children, who are more often cared for at home. Low social support of the mother, use of a child care center, and being an immigrant were risk factors for infections.

No relationship between chronic poverty and growth delay of children was observed in this study. An association often has been reported between family SES and children’s height even in industrialized countries. Our observation of the absence of an association could be the result of a secular tendency toward greater height, which is still ongoing in certain populations. In the Quebec population, children often are observed to be taller than their parents. This type of phenomenon is thought to be especially prevalent among poor individuals because they are of a shorter stature to begin with. As a result, any association between poverty and growth retardation could be confounded by the intergenera-

### Table 1

<table>
<thead>
<tr>
<th>Health Problem</th>
<th>No. of Episodes of Poverty</th>
<th>Total (N = 1950), % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never Poor (n = 1297),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>% (n)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1–2 (n = 268), % (n)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3–4 (n = 280), % (n)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Missing (n = 105), % (n)</td>
<td></td>
</tr>
<tr>
<td>Infections, previous 3 mo*</td>
<td>65.9 (855)</td>
<td>64.6 (173)</td>
</tr>
<tr>
<td></td>
<td>66.8 (187)</td>
<td>63.8 (67)</td>
</tr>
<tr>
<td></td>
<td>65.7 (1282)</td>
<td></td>
</tr>
<tr>
<td>Asthma attacks, previous 12 mo</td>
<td>5.5 (71)</td>
<td>6.0 (16)</td>
</tr>
<tr>
<td></td>
<td>8.9 (25)</td>
<td>5.7 (6)</td>
</tr>
<tr>
<td></td>
<td>6.1 (118)</td>
<td></td>
</tr>
<tr>
<td>Growth retardation*</td>
<td>8.6 (111)</td>
<td>10.4 (28)</td>
</tr>
<tr>
<td></td>
<td>10.4 (29)</td>
<td>14.3 (15)</td>
</tr>
<tr>
<td></td>
<td>9.4 (183)</td>
<td></td>
</tr>
<tr>
<td>Maternal perception of child’s health as less than very good</td>
<td>7.7 (100)</td>
<td>16.8 (45)</td>
</tr>
<tr>
<td></td>
<td>14.3 (40)</td>
<td>12.4 (13)</td>
</tr>
<tr>
<td></td>
<td>10.2 (198)</td>
<td></td>
</tr>
</tbody>
</table>

*Orts, respiratory infections, gastroenteritis, and other infections.

b Height z scores <10th percentile according to the 2000 Centers for Disease Control and Prevention growth charts.

c Infections + asthma + growth retardation.

d Chronic poverty is also associated with a higher score on the CHPI, which means that poor children are more likely to have >1 of the health problems included in the index. Similar results were observed in other studies that used a similar index. This higher number of health problems among poor children might reflect a higher general vulnerability that could result from chronic stress. Stress that results from these difficult living conditions and the biological stress process that ensues can affect the body’s capacity to prevent onset of disease processes.
tional secular trend in height. It also is possible that the poverty measure (below the LICO) that was used in this study is not specific enough, because a previous set of analyses on the same cohort of QLSCD children revealed a relationship between poverty as measured with a scale of a lack of money for basic needs and a growth delay when the children were 2½ and 5 years of age.58 That is, a measure of poverty on the basis of annual household income (as used here) does not provide information on other possible financial resources from members of the extended family or from personal savings. Moreover, with such a measure of poverty, those who are just

<table>
<thead>
<tr>
<th>Control Variable</th>
<th>Never Poor (n = 1297), % (n)</th>
<th>1–2 (n = 268), % (n)</th>
<th>3–4 (n = 280), % (n)</th>
<th>Missing (n = 105), % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, mo</td>
<td>40</td>
<td>45.6 (592)</td>
<td>44.0 (118)</td>
<td>42.9 (120)</td>
</tr>
<tr>
<td></td>
<td>41</td>
<td>50.7 (658)</td>
<td>51.9 (139)</td>
<td>52.9 (148)</td>
</tr>
<tr>
<td></td>
<td>42–44</td>
<td>3.6 (47)</td>
<td>4.1 (11)</td>
<td>4.3 (12)</td>
</tr>
<tr>
<td>Male gender</td>
<td>49.9 (647)</td>
<td>49.6 (133)</td>
<td>49.6 (139)</td>
<td>57.1 (60)</td>
</tr>
<tr>
<td>Birth order</td>
<td>First</td>
<td>46.0 (596)</td>
<td>44.0 (118)</td>
<td>40.7 (114)</td>
</tr>
<tr>
<td></td>
<td>Second-born</td>
<td>40.2 (521)</td>
<td>38.8 (104)</td>
<td>36.8 (103)</td>
</tr>
<tr>
<td></td>
<td>Third-born or more</td>
<td>13.9 (180)</td>
<td>17.2 (46)</td>
<td>22.5 (63)</td>
</tr>
<tr>
<td>Gestational age</td>
<td>&lt;37 wk</td>
<td>4.4 (57)</td>
<td>4.9 (13)</td>
<td>6.1 (17)</td>
</tr>
<tr>
<td>Small for gestational age</td>
<td>7.4 (96)</td>
<td>11.9 (32)</td>
<td>6.1 (17)</td>
<td>9.5 (10)</td>
</tr>
<tr>
<td>Congenital malformation</td>
<td>7.8 (101)</td>
<td>9.7 (26)</td>
<td>7.1 (20)</td>
<td>4.8 (5)</td>
</tr>
<tr>
<td>Duration of breast feeding, wk</td>
<td>Never</td>
<td>24.5 (318)</td>
<td>35.8 (96)</td>
<td>37.5 (105)</td>
</tr>
<tr>
<td></td>
<td>&lt;10</td>
<td>24.3 (315)</td>
<td>23.1 (62)</td>
<td>20.7 (58)</td>
</tr>
<tr>
<td></td>
<td>≥20</td>
<td>36.9 (479)</td>
<td>28.0 (75)</td>
<td>31.4 (88)</td>
</tr>
<tr>
<td>Mother’s characteristics</td>
<td>Age, y</td>
<td>0.6 (8)</td>
<td>4.9 (13)</td>
<td>11.1 (31)</td>
</tr>
<tr>
<td></td>
<td>&lt;20</td>
<td>68.0 (1115)</td>
<td>85.4 (229)</td>
<td>73.6 (206)</td>
</tr>
<tr>
<td></td>
<td>≥35</td>
<td>13.3 (173)</td>
<td>9.7 (26)</td>
<td>15.4 (43)</td>
</tr>
<tr>
<td>Height, percentile</td>
<td>&gt;66th</td>
<td>34.5 (448)</td>
<td>32.1 (86)</td>
<td>28.9 (81)</td>
</tr>
<tr>
<td></td>
<td>33rd–66th</td>
<td>41.5 (538)</td>
<td>38.4 (103)</td>
<td>38.2 (107)</td>
</tr>
<tr>
<td></td>
<td>&lt;66th</td>
<td>23.6 (306)</td>
<td>29.1 (78)</td>
<td>29.3 (82)</td>
</tr>
<tr>
<td>Education</td>
<td>University and undergraduate</td>
<td>73.5 (953)</td>
<td>49.3 (132)</td>
<td>36.1 (101)</td>
</tr>
<tr>
<td></td>
<td>High school diploma</td>
<td>18.7 (242)</td>
<td>28.7 (77)</td>
<td>22.1 (62)</td>
</tr>
<tr>
<td></td>
<td>No high school diploma</td>
<td>7.8 (101)</td>
<td>22.0 (59)</td>
<td>41.4 (116)</td>
</tr>
<tr>
<td>Social-support level</td>
<td>High</td>
<td>69.2 (897)</td>
<td>56.7 (152)</td>
<td>48.6 (136)</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>22.7 (295)</td>
<td>28.0 (75)</td>
<td>26.8 (75)</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>8.1 (105)</td>
<td>15.3 (41)</td>
<td>24.3 (68)</td>
</tr>
<tr>
<td>Physical violence</td>
<td>16.6 (219)</td>
<td>27.6 (74)</td>
<td>31.4 (88)</td>
<td>20 (21)</td>
</tr>
<tr>
<td>Immigrant status</td>
<td>Nonimmigrant/European immigrant</td>
<td>96.7 (1254)</td>
<td>91.8 (246)</td>
<td>77.9 (218)</td>
</tr>
<tr>
<td></td>
<td>Non-European immigrant</td>
<td>3.2 (42)</td>
<td>8.2 (22)</td>
<td>21.8 (61)</td>
</tr>
<tr>
<td>Environment</td>
<td>Type of family (at survey)</td>
<td>86.0 (1116)</td>
<td>61.6 (165)</td>
<td>47.9 (134)</td>
</tr>
<tr>
<td></td>
<td>Intact</td>
<td>9.1 (118)</td>
<td>18.3 (49)</td>
<td>15.4 (43)</td>
</tr>
<tr>
<td></td>
<td>1 nonbiological parent</td>
<td>4.9 (63)</td>
<td>19.8 (53)</td>
<td>36.4 (102)</td>
</tr>
<tr>
<td></td>
<td>Single-parent (mother)</td>
<td>81.0 (1050)</td>
<td>66.0 (177)</td>
<td>57.1 (160)</td>
</tr>
<tr>
<td></td>
<td>Exposure to smoke at home</td>
<td>18.6 (241)</td>
<td>33.6 (90)</td>
<td>41.1 (115)</td>
</tr>
<tr>
<td></td>
<td>None/occasionally</td>
<td>30.1 (391)</td>
<td>41.4 (111)</td>
<td>56.8 (159)</td>
</tr>
<tr>
<td></td>
<td>Daily</td>
<td>39.6 (513)</td>
<td>32.5 (87)</td>
<td>13.9 (39)</td>
</tr>
<tr>
<td>Child care</td>
<td>At home</td>
<td>30.3 (393)</td>
<td>26.1 (70)</td>
<td>29.3 (82)</td>
</tr>
</tbody>
</table>
above the poverty threshold are classified as having sufficient income even though they experience hardships as a result of limited financial resources, which is not the case of those who have a much higher income. This lack of specificity in the poverty measurement also could explain the lack of association between chronic poverty and some of our other health indicators, such as infections.

Our study has other limitations. Although the child’s height was measured by the interviewer and the birth conditions were extracted from hospital records, other information about the child’s health are reported by the mother. We observed in this cohort that maternal perception of child’s health was linked with the presence of health problems. It is also possible that effect sizes for poverty in our modeling are underestimated as a result of overadjustment, because many of the variables that we used as confounders are actually conceptually related to our exposure of interest.

There also are strengths in our study. This birth cohort of children was followed up annually since they were 5 months of age, and the participation rate is high. Data are collected at home by experienced interviewers using questions and instruments that have been used extensively in Canadian and American surveys. Hospital birth records were used to code for birth conditions such as preterm birth, being small for gestational age, and congenital malformations that are known risk factors for health problems during childhood, especially among poor children.

**CONCLUSIONS**

Despite governmental promises to end child poverty, chronic poverty affects a large proportion of families with young children in Quebec. Chronic poverty has important negative consequences for young children’s health beyond being born preterm or small for gestational age and beyond low maternal education. The strength of the association between chronic poverty and health may vary according to different health indicators because not all health problems of preschool children were linked with chronic poverty. Effects of chronic poverty also may vary with the age of the child. Although some health problems are more frequent among young poor children, for other problems, a relationship might emerge later.

It is difficult to generalize the results of this investigation to children from the United States, where access to health care and social policies are different. Moreover, the population studied here is not as heterogeneous as the population in the United States. Remaining disparities in child health despite a universal health care system and a low-cost child care system suggests that unfavorable living conditions and chronic stress that emanates from chronic poverty have a negative impact on child health; therefore, eliminating child poverty should be the target for change.

**ACKNOWLEDGMENTS**

This study was funded by Canadian Institute of Health Research grant 00309MOP-123079. Data were collected...
by the Institut de la Statistique du Québec, Direction Santé Québec.

REFERENCES

1. Phipps S. The Impact of Poverty on Health. Ottawa, Ontario, Canada: Canadian Population Health Initiative and Canadian Institute for Health Information; 2003
4. Phipps S. Rempérissances de la pauvreté sur la santé. Aperçu de la recherche [The Impact of Poverty on Health: A Scan of Research Literature]. Ottawa, Ontario, Canada: Canadian Institute for Health Information; 2003
29. Giles P. Low Income Measurement in Canada. Ottawa, Ontario, Canada: Minister of Industry, Canada; 2004
42. Lin M, Chen Y, Villeneuve PJ, et al. Gaseous air pollutants and asthma hospitalization of children with low household income
54. Monette S, Sèguin L, Gauvin L, Nikiema B. Validation of a measure of maternal perception of the child’s health status. Child Care Health Dev. 2006; In press
Randomized Clinical Trial of Prevention of Hydrocephalus After Intraventricular Hemorrhage in Preterm Infants: Brain-Washing Versus Tapping Fluid

Andrew Whitelaw, MD a,b, David Evans, MD b, Michael Carter, MD c, Marianne Thoresen, MD d, Jolanta Wroblewska, MD d, Marek Mandera, MD f, Janusz Swietlinski, MD e, Judith Simpson, MD g, Constantinos Hajivassiliou, MD h, Linda P. Hunt, PhD d, Ian Pople, MD d

Department of Clinical Science at a North Bristol and d South Bristol, University of Bristol, Bristol, United Kingdom; b Neonatal Intensive Care Unit, Southmead Hospital, Bristol, United Kingdom; c Department of Neurosurgery, Frenchay Hospital, Bristol, United Kingdom; e Department of Neonatal Intensive and Special Care and f Division of Pediatric Neurosurgery, Medical University of Silesia, Katowice, Poland; g Neonatal Intensive Care Unit, Queen Mother’s Hospital, Glasgow, United Kingdom; h Department of Surgical Paediatrics, Royal Hospital for Sick Children, Glasgow, United Kingdom

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

ABSTRACT

OBJECTIVE. Hydrocephalus is a serious complication of intraventricular hemorrhage in preterm infants, with adverse consequences from permanent ventriculoperitoneal shunt dependence. The development of hydrocephalus takes several weeks, but no clinical intervention has been shown to reduce shunt surgery in such infants. The aim of this study was to test a new treatment intended to prevent hydrocephalus and shunt dependence after intraventricular hemorrhage.

METHODS. We randomly assigned 70 preterm infants who had gestational ages of 24 to 34 weeks and were progressively enlarging their cerebral ventricles after intraventricular hemorrhage to either (1) drainage, irrigation, and fibrinolytic therapy to wash out blood and cytokines or (2) tapping of cerebrospinal fluid by reservoir as required to control excessive expansion and signs of pressure (standard treatment). We evaluated outcomes at 6 months of age or hospital discharge (if later).

RESULTS. Of 34 infants who were assigned to drainage, irrigation, and fibrinolytic therapy, 2 died and 13 underwent shunt surgery (dead or shunt: 44%). Of 36 infants who were assigned to standard therapy, 5 died and 14 underwent shunt surgery (dead or shunt: 50%). This difference was not significant. Twelve (35%) of 34 infants who received drainage, irrigation, and fibrinolytic therapy had secondary intraventricular hemorrhage compared with 3 (8%) of 36 in the standard group. Secondary intraventricular hemorrhage was associated with an increased risk for subsequent shunt surgery and more blood transfusions.

CONCLUSIONS. Despite its logical basis and encouraging pilot data, drainage, irrigation, and fibrinolytic therapy did not reduce shunt surgery or death when tested in a multicenter, randomized trial. Secondary intraventricular hemorrhage is a major factor that counteracts any possible therapeutic effect from washing out old blood.
HEMORRHAGE INTO THE VENTRICLES OF THE BRAIN IS ONE OF THE MOST SERIOUS COMPLICATIONS OF PRETERM BIRTH DESPITE IMPROVEMENTS IN THE SURVIVAL OF PRETERM INFANTS. LARGE INTRAVENTRICULAR HEMORRHAGE (IVH) HAS A HIGH RISK FOR NEUROLOGIC DISABILITY, AND >50% OF THESE CHILDREN GO ON TO DEVELOP PROGRESSIVE VENTRICULAR DILATION.\(^1\) INCREASING SURVIVAL OF EXTREMELY PRETERM INFANTS IS ASSOCIATED WITH POSTHEMORRHAGIC VENTRICULAR DILATION (PHVD) WITH HIGH MORBIDITY AND CONSIDERABLE MORTALITY.\(^2\) \(^3\) MULTIPLE BLOOD CLOTS MAY OBSTRUCT THE VENTRICULAR SYSTEM OR CHANNELS OF CEREBROSPINAL FLUID (CSF) REABSORPTION INITIALLY BUT LEAD TO A CHRONIC ARACHNOIDITIS OF THE BASAL CISTERN INVOLVING DEPOSITION OF EXTRACELLULAR MATRIX PROTEINS IN THE FORAMINA OF THE FOURTH VENTRICLE AND THE SUBARACHNOID SPACE.\(^4\) \(^5\) THERE IS EVIDENCE THAT TRANSFORMING GROWTH FACTOR \(\beta\) IS LIKELY TO BE A MEDIATOR OF THIS PROCESS, WHICH PROBABLY TAKES WEEKS,\(^6\) ALTHOUGH THIS WAS NOT CONFIRMED IN 1 STUDY THAT FOUND THAT VASCULAR ENDOTHELIAL GROWTH FACTOR WAS ELEVATED IN THE CSF OF INFANTS WITH PHVD.\(^6\) INTRAVENTRICULAR BLOOD AND VENTRICULAR EXPANSION MAY HAVE ADVERSE EFFECTS ON THE IMMATURE PERIVENTRICULAR WHITE MATTER BY A VARIETY OF MECHANISMS, INCLUDING PHYSICAL DISTORTION, RAISED INTRACRANIAL PRESSURE (ICP),\(^7\) FREE RADICAL GENERATION FACILITATED BY FREE IRON,\(^8\) AND INFLAMMATION.\(^9\)

TREATMENT IS MORE DIFFICULT THAN OTHER TYPES OF HYDROCEPHALUS, BECAUSE THE LARGE AMOUNT OF BLOOD AND PROTEIN LEVEL IN THE CSF COMBINED WITH THE SMALL SIZE AND INSTABILITY OF THE PATIENT MAKES AN EARLY VENTRICULOPERITONEAL SHUNT OPERATION IMPOSSIBLE. THERE IS A CONSIDERABLE COMPLICATION RATE THROUGHOUT THE CHILD’S LIFE FROM SUCH SURGERY, AND THE CHILD IS PERMANENTLY DEPENDENT ON THE SHUNT SYSTEM.\(^10\) NEITHER TREATMENT BY REPEATED LUMBAR OR VENTRICULAR TAPPING NOR THE USE OF ACETAZOLAMIDE AND FUROSEMIDE TO REDUCE CSF PRODUCTION PREVENTS THE NEED FOR SHUNT SURGERY OR IMPROVES NEUROLOGIC OUTCOME, AND BOTH HAVE APPRECIABLE ADVERSE EFFECTS.\(^11\) \(^12\) DEATH OR SHUNT SURGERY OCCURRED IN 67% AND 52% OF INFANTS, RESPECTIVELY, IN THESE 2 TRIALS.\(^11\) \(^12\) PHASE 1 CLINICAL TRIALS OF INTRAVENTRICULAR FIBRINOLYTIC THERAPY TREATMENT\(^13\) \(^14\) AND A SMALL RANDOMIZED TRIAL\(^15\) HAVE NOT GIVEN ENCOURAGING RESULTS.

THE STANDARD TREATMENT VARIATES, AND FEW CENTERS HAVE BUILT UP LARGE SERIES FOR ANALYSIS. THE STANDARD Arms OF THE VENTRICULOMEGALY TRIAL AND THE PHVD DRUG TRIAL BOTH USED SELECTIVE TAPPING OF CSF TO CONTROL SIGNS OF PRESSURE OR EXCESSIVE HEAD ENLARGEMENT.\(^11\) \(^12\) THE PRACTICE OF INSERTING A VENTRICULAR ACCESS DEVICE, SUCH AS AN OMMAYA RESERVOIR, TO FACILITATE REPEATED TAPPING OF ADEQUATE VOLUMES IS WIDELY PRACTICED WITHOUT HAVING BEEN TESTED BY RANDOMIZED TRIAL.\(^16\)

ADULTS WITH IVH HAVE BEEN TREATED BY EARLY VENTRICULAR DRAINAGE COMBINED WITH INTRAVENTRICULAR RECOMBINANT TISSUE PLASMINOGEN ACTIVATOR (rtPA), AND THESE STUDIES REPORTED MORTALITY THAT WAS MUCH LOWER THAN HISTORICAL CONTROLS.\(^17\) \(^18\) CLEARLY, IVH IN ADULTS IS VERY DIFFERENT FROM IVH IN PRETERM INFANTS WITH RESPECT TO ETIOLOGY AND ALSO LIKELIHOOD OF RAISED ICP, BUT THE ADULT EXPERIENCE AND OUR PREVIOUS RESEARCH ON THIS TOPIC SUGGESTED THAT THERE MIGHT BE A CRITICAL PERIOD OF DAYS OR WEEKS DURING WHICH REMOVING AS MUCH BLOOD, CYTOKINE, AND FREE IRON AS POSSIBLE IN PRETERM INFANTS WITH LARGE IVH MIGHT STOP THE PROGRESSION OF PATHOLOGY BEFORE IRREVERSIBLE HYDROCEPHALUS IS ESTABLISHED.

WE HAVE PILOTED A PROCEDURE, DRAINAGE, IRRIGATION, AND FIBRINOLYTIC THERAPY (DRIFT), THAT AIDS TO REMOVE INTRAVENTRICULAR BLOOD AND THE CYTOKINES THAT ARE ASSOCIATED WITH HYDROCEPHALUS BEFORE HYDROCEPHALUS BECOMES ESTABLISHED.\(^19\) TWENTY-FIVE PRETERM INFANTS WITH VENTRICULAR ENLARGEMENT AFTER A LARGE IVH WERE ENROLLED. EIGHTEEN (75%) OF 24 SURVIVORS DID NOT REQUIRE A VENTRICULOPEITONEAL SHUNT. TWO INFANTS DEVELOPED RESERVOIR-ASSOCIATED INFECTION, AND 2 INFANTS HAD A SECOND IVH. SHUNT SURGERY OR DEATH (28%) WAS REDUCED COMPARED WITH HISTORICAL CONTROLS WITH SIMILAR TREATMENT CRITERIA.\(^11\) \(^12\) ALTHOUGH THIS TECHNIQUE IRRIGATES AND DRAINS THE LATERAL VENTRICLES, WE HAD ULTRASOUND EVIDENCE THAT BLOOD CLOT WAS ALSO REMOVED FROM THE THIRD VENTRICLE. THE PRIMARY OBJECT OF THIS STUDY WAS TO TEST THE HYPOTHESIS THAT TREATMENT BY DRIFT REDUCES VENTRICULOPEITONEAL SHUNT OR DEATH WHEN COMPARED WITH STANDARD TREATMENT FOR PHVD.

METHODS

Study Infants

The study was approved by the research ethics board of each institution that took part: Southmead Hospital (Bristol, United Kingdom), Royal Hospital for Sick Children (Glasgow, United Kingdom), Medical University of Silesia (Katowice, Poland), and Haukeland Hospital (Bergen, Norway). Written informed consent was obtained from the mother of each infant. In all 4 centers, preterm infants who required intensive care or showed neurologic abnormalities had daily cranial ultrasound scans for the first 3 days and then at least weekly for 4 weeks. When IVH was diagnosed, ultrasound scanning was then done twice weekly. Ventricular measurements were made when there was any visible enlargement, and head circumference was then measured daily. Cranial ultrasound continued twice weekly until resolution of ventricular enlargement and more frequently when enlargement was progressive.

Inclusion criteria were (1) IVH documented on ultrasound; (2) age no more than 28 days; and (3) progressive dilation of both lateral ventricles with each side: (a) ventricular width 4 mm over the 97th centile;\(^19\) (b) all of the following: anterior horn diagonal width 4 mm (1 mm over 97th centile);\(^19\) thalamo-occipital distance 26 mm (1 mm over 97th centile);\(^19\) and third ventricle width 3 mm (1 mm over 97th centile); or (c) measurements above a or b on 1 side combined with obvious midline shift indicating a pressure effect. Exclusion cri-
Criteria were a prothrombin time of $>20$ seconds or accelerated partial thromboplastin time of $>50$ seconds or platelets $<50\,000/mL$.

**Randomization**

A computer-generated randomization scheme was used to assign the infants to treatment groups in a 1:1 ratio. Randomization was stratified by study center and in blocks of 8, 10, or 12. Each infant was allocated to treatment using sequentially numbered, double-opaque envelopes (1 envelope inside the other for security) that each contained a “DRIFT” or “standard treatment” card.

**Treatment**

**DRIFT**

A detailed description of this technique was published previously. Under anesthesia, 2 ventricular catheters were inserted (right frontal and left occipital). A Codman external ventricular drainage system (Johnson & Johnson, Piscataway, NJ) was connected. The connections are shown in Fig 1. rTPA (Actilyse, Boehringer Ingelheim Int, Ingelheim, Germany) 0.5 mg/kg was injected intraventricularly. After 8 hours, artificial CSF (Torbay Pharmaceutical Manufacturing Unit, Kemmings Close, Paignton, United Kingdom; and the Biochafa Production Plant, Sosnowiec, Poland) with 10 mg of vancomycin and 5 mg/500 mL intrathecal gentamicin was infused at 20 mL/hour into the right frontal ventricular catheter with a pressure transducer on the in-going line. Simultaneously, fluid was allowed to drain from the left occipital ventricular catheter, the height of the drainage reservoir being raised or lowered to maintain the ICP below 7 mm Hg. It was usually necessary to drain 60 to 100 mL/24 hours more than the infused volume to maintain normal pressure. The drainage fluid was usually cola-colored initially and then gradually cleared. When the drainage fluid became colorless, infusion of artificial CSF was stopped and both catheters were removed after a median of 3 days (range: 2–7 days).

**Standard Treatment**

The infant was observed daily for raised ICP (irritability, apnea, reduced consciousness, bulging fontanelle, or loss of diastolic velocities on cerebral arteries) or excessive head enlargement over time. When neither of these applied, no intervention was conducted. This policy followed logically from the previous randomized trial data on PHVD. Head circumference enlarges by $\sim 1\,\text{mm/day}$ between 26 weeks of gestation and 32 weeks and by $\sim 0.7\,\text{mm/day}$ between 32 and 40 weeks. Excessive head enlargement was defined as 2 mm/day. Suspected raised ICP or excessive head expansion prompted a lumbar puncture (LP) with the object of removing 10 to 20 mL/kg over 10 to 20 minutes. When the LP was successful, the infant continued to be observed to determine whether a repeat LP would be necessary. When the head circumference increased by 2 mm/day after LP, additional tapping was conducted.

**Criteria for Insertion of Ventricular Reservoir**

When $>2$ LPs were necessary or when the LP failed to drain enough to normalize head growth to $<2\,\text{mm/day}$, a ventricular reservoir was indicated. When there was a delay in surgical insertion of a reservoir and the clinical situation seemed urgent, a ventricular tap was conducted to drain 10 to 20 mL/kg over 10 to 20 minutes at a frequency to ensure that head growth was $<2\,\text{mm/day}$, preferably 1 mm/day.

**Failed DRIFT and Crossover to Standard Treatment**

When DRIFT was followed by persistent enlargement of ventricles and excessive head growth (2 mm/day), management changed to standard treatment with LPs and ventricular reservoir. Infants were not switched from conventional therapy to DRIFT.

**Criteria for Shunt Surgery**

When an infant was having repeated reservoir taps to control head growth, this was continued until weight reached 2500 g and CSF protein fell to $<1.5\,\text{g/L}$. Tapping was stopped and the head circumference was measured daily. When the head circumference increased by 2 mm/
day, ultrasound was used to confirm that expansion was CSF and not brain growth. A ventriculoperitoneal shunt was indicated when the excessive growth persisted over several days after the described procedures.

Outcomes
In this study, follow-up was until 6 months of age or discharge from hospital, when later. The primary outcome is a composite of ventriculoperitoneal shunt surgery and death. This was documented from the hospital clinical records and by our own contact with the family.

Secondary IVH was diagnosed on the basis of ultrasound appearance of new intraventricular echodensities (Fig 2) within 1 week of randomization combined with a fall in hemoglobin of at least 2 g/dL in 2 days. The ultrasound diagnosis was always confirmed by at least 2 examiners. Diagnosis of secondary CSF infection required a positive culture of bacteria and a raised white cell count in the CSF after randomization. Neurodevelopmental outcome at 2 years past term will be the subject of a separate report.

Statistical Analysis
On the basis of previous trials, the inclusion criteria predicted that 50% to 60% of the infants would need a shunt operation or die. Our initial power calculation (using a 5% level of significance) indicated that 60 infants in each group would give 91% power of detecting

FIGURE 2
Secondary IVH. A, Coronal ultrasound view with dilation of both lateral and the third ventricles and clot (white) in the right ventricle but not the left. B, Parasagittal view of the left ventricle with only very small clots visible. C, Corresponding coronal view after secondary hemorrhage during DRIFT. The ventricles are smaller because of the drainage, but there is now a blood clot in the left ventricle. D, The whole left ventricle is now full of blood clot.
a reduction from 60% in primary outcome to 30% and a 79% power of detecting a reduction from 55% to 30%.21

The trial was closed early after an interim analysis of the first 50% of infants (30 DRIFT and 30 standard treatment) conducted for the Data Monitoring and Safety Group. “Conditional” power was calculated for the primary outcome (ie, the probability of obtaining a significant difference in the rate of shunt/death given the results that had been obtained up to the half-way point).22,23 The estimate of power under the alternative trial hypothesis (30% DRIFT versus 55% standard) was only 15%. The Data Monitoring and Safety Group made a recommendation to the Trial Steering Group in March 2006 to cease recruitment, by which time an additional 10 infants had been treated. The statistical analysis presented here is for all 70 infants.

Analysis was by intention to treat. Standard statistical tests were used throughout. Proportions were compared using continuity-corrected χ² tests, a Student’s t test was used to compare the mean minimum hemoglobin concentration, and a Mann-Whitney U test was used to compare the number of transfusions.

RESULTS
Recruitment started in Bristol in February 2003 and finished in April 2006 (38 months). Seventy infants were recruited to the DRIFT trial: 47 in Bristol, 20 in Katowice, 2 in Glasgow, and 1 in Bergen. The great majority of infants who were recruited in Bristol and Katowice were transferred from other cities for neurological assessment. Seventy of 74 parents of the infants who met trial criteria gave consent when asked, and none of the recruited infants was lost to follow-up. Patient flow is shown in Fig 3.

Table 1 shows the characteristics of the infants at randomization. The 2 treatment groups were similar with respect to gender, gestation, birth weight, age at randomization, and presence of parenchymal hemorrhagic infarction at entry.

Table 2 shows the main outcomes at discharge. Of 34 infants who were assigned to DRIFT, 2 died and 13 underwent shunt surgery (dead or shunt: 44%). Of 36 infants who were assigned to standard therapy, 5 died (1 with a shunt) and 13 survived shunt operation (dead or shunt: 50%). This difference was not significant (P = .80). The causes of death were not directly related to
brain injury or neurosurgical intervention but were necrotizing enterocolitis, septicemia, and chronic lung disease.

Infants who required shunt surgery in the DRIFT group were slightly younger when they received this than the corresponding infants in the standard treatment group (median [range]: 72 [40–221] vs 99 [42–147] days; \( P = .25 \)).

Infants in the standard group were almost twice as likely to have a ventricular reservoir inserted (38% vs 75%; \( P = .004 \)). Within the standard group, 15 (56%) of 27 infants who received a reservoir subsequently required a shunt. None of the infants in the standard group without a reservoir needed a shunt. In the DRIFT group, 6 of 21 infants who did not receive a reservoir eventually needed a shunt. They avoided a reservoir because of repeated LPs or because enlargement was late enough for shunt surgery to be possible.

Secondary IVH was an important secondary outcome because this occurred in the pilot study of DRIFT. Twelve (35%) of 34 infants who received DRIFT had secondary IVH, compared with 3 (8%) of 36 in the standard treatment group, and this difference was statistically significant (difference 27%; 95% confidence interval: 9% to 45%; \( P = .014 \)). It was not possible to time secondary IVH precisely because infants were usually scanned once a day, but most occurred within 24 hours of catheter insertion. Because secondary IVH was accompanied by a fall in hemoglobin, there often was a requirement for replacement blood transfusion. The mean number of blood transfusions that were received in the first 7 days after randomization was 1.7 (range: 0–4) in the DRIFT group and 0.8 (range: 0–2) in the standard group (\( P = .001 \)). Within the DRIFT group, the mean number received was 2.2 (range: 1–4) when there was secondary IVH and 1.4 (range: 0–4) if there was no secondary IVH (\( P = .055 \)). The mean of the lowest hemoglobin was 8.9 (SD: 1.3) g/dL in the DRIFT group and 9.7 (SD: 1.3) g/dL in the standard group (\( P = .012 \)). In only 1 case was the secondary IVH clinically apparent. This infant was undergoing DRIFT when he acutely developed thrombocytopenia. There was no significant difference between centers in the frequency of secondary IVH.

Additional evidence that secondary IVH is of clinical importance is provided by 8 (67%) of 12 infants in the DRIFT group who had secondary IVH that required shunt surgery, whereas only 5 (23%) of 22 in the DRIFT group without secondary IVH required shunt surgery (\( P = .032 \)). Table 3 shows that the infants who bled were not different from those who did not bleed with respect

| TABLE 1 Characteristics of Infants When Randomly Assigned Into the Trial |
|----------------|----------------|----------------|
| Characteristic | DRIFT (N = 34) | Standard (N = 36) |
| Center         |                |                |
| Bristol       | 22             | 25             |
| Katowice      | 10             | 10             |
| Glasgow       | 1              | 1              |
| Bergen        | 1              | 0              |
| Gestation, wk | 27 (24–34)     | 28 (24–35)     |
| Birth weight, g | 1066 (640–2100) | 1079 (720–2755) |
| Male, n (%)   | 24 (71)        | 23 (64)        |
| Parenchymal hemorrhagic infarction, n (%) | 18 (53) | 18 (50) |
| Age at randomization, d | 20 (7–28) | 18 (9–28) |
| Ventricular width, mm | 16 (12–21) | 18 (12–30) |

| TABLE 2 Trial Outcomes at 6 Months or First Discharge Home (When Later) |
|----------------|----------------|----------------|
| Outcome                  | DRIFT (N = 34), n (%) | Standard Treatment (N = 36), n (%) | Difference (DRIFT – Standard), % (95% CI) | Relative Risk (95% CI) |
| VP shunt                | 13 (38)         | 14 (39; 1 dead) | -1 (-23 to 22) | 0.98 (0.54 to 1.78) |
| Dead                    | 2 (6)           | 5 (14)         | -8 (-22 to 6)  | 0.42 (0.09 to 2.04) |
| Dead or shunt           | 15 (44)         | 18 (50)        | -6 (-29 to 17) | 0.88 (0.54 to 1.45) |
| Reservoir               | 13 (38)         | 27 (75)        | -37 (-58 to -15) | 0.51 (0.32 to 0.81) |
| Secondary IVH           | 12 (35)         | 3 (8)          | 27 (9 to 45)   | 4.24 (1.31 to 13.72) |
| Secondary infection     | 0 (0)           | 1 (3)          |                |                |
| Dead or shunt, by center |                |                |                |                |
| Bristol                 | 8/22 (36)       | 14/25 (56)     | -20 (-48 to 8) | 0.65 (0.34 to 1.25) |
| Katowice                | 6/10 (60)       | 4/10 (40)      | 20 (-23 to 63) | 1.5 (0.60 to 3.74) |
| Glasgow                 | 1/1             | 0/1            |                |                |
| Bergen                  | 0/1             |                |                |                |

CI indicates confidence interval.
to age at randomization, initial coagulation status, presence of parenchymal hemorrhagic infarction, gestational age, birth weight, or gender. Secondary infection in the CSF was rare, with none in the DRIFT group and 1 in 36 of the standard group (coagulase-negative *Staphylococci* in an infant with a reservoir), which resolved with standard intravenous antibiotic therapy.

**DISCUSSION**

This trial arose from the negative results of previous trials of therapies for posthemorrhagic ventricular dilation and from progress in understanding the pathophysiology of this condition. Available evidence suggested the possibility that progression of hydrocephalus might be halted by removing as much of the blood and cytokines as possible before irreversible hydrocephalus. Ventricular drainage and irrigation are established techniques, and the rTPA was added to increase mobilization of old blood clot and to reduce catheter blockage. A single-center pilot study of DRIFT had shown promising results with low rates of secondary hemorrhage (<10%), infection (<10%), and shunt surgery or death (28%) when compared with historical controls. There was considerable interest in the pilot study, and we were strongly advised to test DRIFT in a multicenter, randomized, clinical trial.

Much thought had to be given to the inclusion criteria and what constituted “standard treatment” for PHVD because of important variation among centers. Because all invasive interventions for PHVD have the potential to cause infection and hemorrhage, we used entry criteria that have predicted a high rate of shunt surgery and disability. On the basis of the evidence that raised ICP reduces cerebral perfusion and can adversely affect neurophysiology,21–24 our standard therapy involved tapping off fluid when there was suspicion of raised ICP or excessive head enlargement.

Strengths of this trial are that (1) the 2 treatment groups were well matched, (2) no infants were lost to follow-up, and (3) all infants received the allocated treatment. A weakness is the limited number of infants, but this was a consequence of the power calculation, the planned interim analysis, and subsequent intervention of the Data Monitoring and Safety Group.

The trial showed that there was no significant difference between the 2 treatment groups in the primary outcome, shunt surgery or death. Even if recruitment were to continue to 120, the trial was highly unlikely to demonstrate a significant advantage for the DRIFT group. An important finding was the significantly increased secondary IVH in the DRIFT group. Although these events were clinically silent in all but 1 case and revealed only by careful examination of daily ultrasound scans and daily hemoglobin measurements, they were associated with increased need for blood transfusion and shunt surgery. Because the central objective of DRIFT as a treatment was to remove as much blood as possible, secondary IVH would directly counteract the planned therapeutic process; therefore, it is understandable that secondary IVH was followed by an increased risk for permanent hydrocephalus. It is not possible from our data to predict which infants are at particularly high risk for secondary IVH.

It must be remembered that, unlike the use of rTPA after clipping an aneurysm with subarachnoid hemorrhage, the original site of intraventricular bleeding in the infants in the DRIFT trial had not been repaired surgically, and it may be that the injection of TPA uncovered the original bleeding site or facilitated bleeding from 1 of the puncture sites that were necessary to place the ventricular catheters. However, we had previously found that omitting rTPA when the CSF was bloody sometimes resulted in the blocking of ventricular catheters with clots.

A multicenter trial of a new technique depends on effective teaching and dissemination. The pilot DRIFT study involved only 1 pediatric neurosurgeon, 2 neonatologists, and a small group of neonatal nurses. In contrast, the multicenter, randomized DRIFT trial involved a much larger number of neurosurgeons, neonatologists, and nurses. Despite that considerable time was devoted to training, there may have been variation in practice and technique that translated into more variable outcomes.

**CONCLUSIONS**

DRIFT did not reduce ventriculoperitoneal shunt surgery or death in preterm infants with ventricular dilation after IVH when compared with tapping of CSF to control excessive head expansion or raised ICP. Tapping a ventricular reservoir was relatively safe and effective in controlling hydrocephalus even in extremely small infants. It may be that earlier use of a reservoir for tapping can

---

**TABLE 3** Comparison of DRIFT Infants With and Without Secondary IVH

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>DRIFT Group Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Secondary IVH ((N = 12))</td>
</tr>
<tr>
<td>Gestation, median (range), wk</td>
<td>27 (24–32)</td>
</tr>
<tr>
<td>Birth weight, median (range), g</td>
<td>1046 (720–1720)</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Parenchymal infarction, n (%)</td>
<td>7 (58)</td>
</tr>
<tr>
<td>Age at randomization, median (range), d</td>
<td>21 (12–28)</td>
</tr>
</tbody>
</table>
improve outcome, and we understand that this is being tested prospectively in the Netherlands. Preventing permanent hydrocephalus in this group of high-risk infants is still inhibited by a limited knowledge of the pathogenesis.

ACKNOWLEDGMENTS

This study was supported by Cerebra and the James and Grace Anderson Trust and North Bristol NHS Trust, which was the sponsor in England.

The following investigators participated in the DRIFT trial: A. Whitelaw, I. Pople, M. Carter, D. Evans, and M. Thoresen (Bristol, United Kingdom); J. Wroblewska, M. Mandera, J. Swietlinski, and E. Musialik-Swietlinska (Katowice, Poland); J. Simpson, A. Watt, R. Carachi, and C. Hajivassilou (Glasgow, United Kingdom); and H. Reigstad (Bergen, Norway); the Trial Steering Group consisted of N. Marlow, A. Whitelaw, I. Pople, D. Evans, J. Wroblewska, J. Simpson, C. Hajivassilou, L. Hunt, S. and R. Walker-Cox (parents of a child in the DRIFT pilot study); the Data Monitoring and Safety Group was made up by A. Wilkinson (Chair), S. Gates, and A. Hockley; and hospitals that referred infants to the DRIFT trial were University College London Hospital, Singleton Hospital (Swansea), Cheltenham General Hospital, Gloucestershire Royal Hospital, Hammersmith Hospital (London), Hillingdon Hospital (London), University Hospital of Wales (Cardiff), Birmingham Maternity Hospital, Dudley Road Hospital (Birmingham), Royal Sussex County Hospital (Brighton), St Mary’s Hospital (Portsmouth), Basildon General Hospital, Royal Gwent Hospital (Newport), Musgrove Park Hospital (Taunton), Royal Exeter and Devon Hospital, and Royal Cornwall Hospital. Infants were referred to Katowice from the hospitals at Zabrze, Bielsko-Biała, Opole, Gdansk, and Lomza.

We are grateful to the nursing and medical staff at the trial centers, especially Chrissie Israel and Sue Lamburne in Bristol and Dr Ewa Musialik-Swietlinska and Prof Florian Ryszka in Katowice, and the hospitals that referred infants to the trial centers. We thank the parents of all of the infants in the DRIFT trial for their trust and commitment.

REFERENCES

Neurodevelopmental and Growth Outcomes of Extremely Low Birth Weight Infants Who Are Transferred From Neonatal Intensive Care Units to Level I or II Nurseries

Shabnam Lainwala, PhD, MBBSa, Rebecca Perritt, MSb, Kenneth Poole, PhDb, Betty Vohr, MDb, for the National Institute of Child Health and Human Development Neonatal Research Network

*Department of Pediatrics, Women and Infants Hospital, Providence, Rhode Island; †Statistical and Epidemiology Unit, RTI International, Research Triangle Park, North Carolina

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

ABSTRACT

OBJECTIVE. Transfer of clinically stable infants to level I and II nurseries alleviates demands on NICUs and allows better use of beds and resources. This study compared growth, neurodevelopmental impairments, postdischarge rehospitalization and deaths, and compliance for follow-up assessment at 18 to 22 months’ corrected age of extremely low birth weight infants who transferred to level I and II nurseries with those who continued to receive care to discharge in a NICU.

METHODS. A retrospective analysis of prospectively collected data from the National Institute of Child Health and Human Development Neonatal Research Network was performed. Between January 1998 and June 2002, 4896 infants born with birth weights of 401 to 1000 g and cared for in 19 National Institute of Child Health and Human Development Neonatal Research Network centers were included. The sample consisted of 4392 survivors who received continuing care in the NICU to discharge home and 504 infants who were transferred to level I and II nurseries before discharge home. Demographics, perinatal characteristics, growth, and neurodevelopmental impairments were compared. Bivariate and logistic regression analyses were performed.

RESULTS. Transfer of infants to level I and II nurseries was associated significantly with white race, private insurance, outborn status, and lower neonatal morbidities and compliance for follow-up compared with the NICU group. After adjusting for known covariates, transfer to level I and II nurseries was not associated with neurodevelopmental impairments or death; however, it was associated with increased postdischarge rehospitalization.

CONCLUSIONS. Extremely low birth weight infants who are transferred to level I and II nurseries have similar growth and neurodevelopmental outcomes to infants who are discharged from a NICU. They are, however, more likely to be readmitted to the hospital and are less compliant for follow-up. Establishment of consistent guidelines for comprehensive discharge planning for level I and II nurseries may improve follow-up compliance and reduce rehospitalization rates among these infants who are transferred.
The driving forces for the emergence of perinatal regionalization in the 1970s were shortage of trained personnel who are required for the care of low birth weight infants and the financial costs of maintaining these skills. This change in the pattern of services provided has improved the morbidity and the mortality of low birth weight, high-risk infants who are born at level I and II nurseries in community hospitals and transferred to a level III tertiary medical center NICU for additional care. Transfer of high-risk mothers from community hospital obstetric units to tertiary perinatal centers before delivery has further improved the morbidity and the mortality in this group of infants.

In the past 3 decades, advancing perinatal and neonatal technology, changing socioeconomic conditions, and health behavior patterns have resulted in an increase in the preterm birth rate, as well as survival, of these high-risk infants who require NICU care. Regional studies of the transport system have shown a 40% increase in transfer of high-risk infants to NICUs from community hospital level I and II nurseries. This has introduced new demands on the resources of tertiary centers, such as increased need for beds, equipment, skilled staff, and finances. Transfer of clinically stable infants to level I and II nurseries alleviates some of these demands on tertiary NICUs as well as allows for better use of level I and II nursery beds and resources.

Although a limited number of studies reported better short-term clinical outcomes of infants who were transferred from a NICU to level I and II nurseries, an extensive literature review did not identify reports on the neurodevelopmental outcomes of these infants. The objective of our study was to compare neonatal and perinatal characteristics and growth and neurodevelopmental outcomes at a follow-up visit at 18 to 22 months’ corrected age (CA) of extremely low birth weight (ELBW) infants who received total care in the tertiary center NICU with infants who were transferred to a community level I or II nursery before discharge home. It was hypothesized that infants who were transferred to level I and II nurseries and infants who continued to receive care in the tertiary center NICU until discharge would have similar neurodevelopmental and growth outcomes at follow-up.

METHODS
This retrospective study was conducted on prospectively collected data on all ELBW infants who were admitted within 14 days of birth between January 1998 and June 2002 with birth weights of 401 to 1000 g at 1 of the 19 level III neonatal centers of the National Institute of Child Health and Human Development (NICHD) Neonatal Network and who survived to discharge home from the NICU or were transferred to level I and II nurseries. Data were obtained from the NICHD Neonatal Network database. Each center’s participation was approved by its respective institutional review board. Research coordinators collected demographic, perinatal, and infant data at each participating center using definitions that were developed by the investigators and described in previous publications. Neonatal outcome data were assessed at 120 days after birth, at discharge, or at the time of death. Hospital data on infants who were transferred to level I and II nurseries were collected until they were discharged from the hospital. Data including network NICU transfer rates and level of care provided at their affiliate level I and II nurseries.

Length of hospitalization of infants in the NICU included the number of days in the tertiary center NICU until discharge to home, whereas length of hospitalization of the level I and II nurseries group included the number of days in the NICU as well as level I and II nurseries until discharge to home. Demographics including maternal age, race, education, insurance type, gravidity, singleton or multiple birth, and infant birth weight and estimated gestational age were obtained. Neonatal morbidities including the incidence of respiratory distress syndrome (RDS), intraventricular hemorrhage (IVH), periventricular leukomalacia (PVL), necrotizing enterocolitis (NEC), retinopathy of prematurity (ROP), early- and/or late-onset sepsis, bronchopulmonary dysplasia (BPD), and need for home oxygen therapy were collected.

IVH was reported according to the classification of Papile et al. PVL was diagnosed by cranial ultrasonography performed at >2 weeks of age. Cranial ultrasonographs were read by individual radiologists at each center and not by a central reader. Early sepsis was defined as sepsis with a positive blood culture drawn at <72 hours of age, and late-onset sepsis was defined as a positive blood culture drawn ≥72 hours of age. NEC stage II or higher was recorded using a modified Bells classification. BPD was defined as a supplemental oxygen requirement at ≥28 days of age.

Follow-up neurologic, developmental, and growth assessments were performed at 18 to 22 months’ CA. The neurodevelopmental tests are valid to 36 months’ CA. These neurodevelopmental tests were age-adjusted for infants who were seen beyond the 18- to 22-months’ CA window. The neurologic examination based on the Amiel-Tison neurologic assessment was conducted by certified examiners. Cerebral palsy was defined as a nonprogressive central nervous system disorder characterized by abnormal muscle tone in at least 1 extremity and abnormal control of movement and posture. The Bayley Scales of Infant Development II mental and motor scales were administered by certified testers. A Bayley score of 100 ± 15 represents the mean ± 1 SD. A score of <70 is 2 SD below the mean. Hearing status information was obtained from parents and follow-up audiologic test results. Hearing impairment was defined as using hearing aids in both ears. An examination of the
eyes was completed, and eye examination history was obtained from parents and from physician reports, when available. Blind was defined as no useful vision (>20:200) in both eyes, when corrected. Neurodevelopmental impairment (NDI) was defined as presence of 1 or more of the following: mental developmental index (MDI) <70, psychomotor developmental index (PDI) <70, cerebral palsy, hearing impairment, or blindness. Growth parameters including weight, length, and head circumference measurements were assessed using standardized techniques. Percentiles were determined using growth charts developed by the National Center for Health Statistics (published May 30, 2000). Immunization information, including respiratory syncytial virus prophylaxis, was obtained from parents.

The 2 study groups included infants who received total care in the tertiary center NICU until discharge home and infants who were transferred from a NICU to level I and II nurseries for continuing care before discharge home. Bivariate analyses of demographic characteristics, neonatal morbidities, growth parameters, neurodevelopmental outcomes at the follow-up at 18 to 22 months’ CA, compliance for follow-up at 18 to 22 months CA, and rehospitalizations and deaths after discharge were performed. Also, infants who were lost to follow-up were compared with those who were followed for neurodevelopmental and growth assessment using bivariate analysis. Numbers and percentages are shown for categorical covariates. Means and SDs are shown for continuous covariates. For comparison of NICU and level I and II nurseries groups, $\chi^2$ tests were performed for categorical covariates, and Kruskal-Wallis nonparametric tests were performed for continuous covariates. Logistic regression analyses was performed to evaluate the effects of transfer to level I and II nurseries and outborn status on NDI, MDI and PDI <70, postdischarge rehospitalization in the first year of life and before the follow-up visit at 18 to 22 months’ CA, and loss to follow-up at 18 to 22 months’ CA.

### RESULTS

A total of 4896 infants with birth weight between 401 and 1000 g were born at the 19 network centers or transferred in within 14 days of birth between January 1998 and June 2002 and survived to be transferred to level I and II nurseries or discharge home; 4392 (89.7%) infants received complete care in the NICU, and 504 (10.3%) infants were transferred to level I and II nurseries before discharge home. Twenty-three infants who were transferred from a network NICU to another level III NICU and 41 infants who were transferred to long-term care facilities were excluded from the study. As expected, there was a significantly high risk for NDI (75%) and death (11%) in the group of infants who were transferred to long-term care facilities.

All network NICUs transferred infants to level II nurseries, and 50% of NICUs transferred infants to level I nurseries. During the study period, the network NICU transfer rates to level I and II nurseries ranged from 1% to 44%. Among the units that received transfers, 80% of level II nurseries were staffed by a neonatologist and 53% of these level II nurseries provided conventional ventilatory care to the infants, whereas only 20% of level I nurseries were staffed by a neonatologist and none provided ventilatory care. None of the level I or level II nurseries had admission criteria for infant gestational age or weight.

The maternal demographic and socioeconomic characteristics of the NICU and level I and II nurseries groups are presented in Table 1. There were no significant differences in maternal age, education, gravidity, and multiple gestation between NICU and level I and II nurseries groups. Infants who were white compared with other ethnic groups ($P < .0001$) and infants with private health insurance ($P = .002$) were more likely to be transferred to level I and II nurseries.

Table 2 shows the neonatal characteristics and morbidities of the NICU and level I and II nurseries groups. Infant birth weight and gestational age were significantly higher in the level I and II nurseries group ($P \leq .0005$) compared with the NICU group. There were no signifi-

### TABLE 1 Maternal Demographic and Socioeconomic Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>NICU, n (%)</th>
<th>Level I and II Nurseries, n (%)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>4392</td>
<td>504</td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td>.1</td>
</tr>
<tr>
<td>&lt;20</td>
<td>691 (16)</td>
<td>96 (19)</td>
<td></td>
</tr>
<tr>
<td>20–34</td>
<td>3067 (70)</td>
<td>333 (66)</td>
<td></td>
</tr>
<tr>
<td>&gt;34</td>
<td>633 (14)</td>
<td>75 (15)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>.0001</td>
</tr>
<tr>
<td>White</td>
<td>1638 (37)</td>
<td>282 (56)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>2018 (46)</td>
<td>111 (22)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>615 (14)</td>
<td>89 (18)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>79 (2)</td>
<td>12 (2)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>41 (1)</td>
<td>9 (2)</td>
<td></td>
</tr>
<tr>
<td>Health insurance*</td>
<td></td>
<td></td>
<td>.002</td>
</tr>
<tr>
<td>Private</td>
<td>1418 (35)</td>
<td>180 (43)</td>
<td></td>
</tr>
<tr>
<td>Public</td>
<td>2651 (65)</td>
<td>243 (57)</td>
<td></td>
</tr>
<tr>
<td>Education less than high school*</td>
<td>1131 (27)</td>
<td>102 (23)</td>
<td>.07</td>
</tr>
<tr>
<td>Gravida &gt;1</td>
<td>2973 (68)</td>
<td>324 (64)</td>
<td>.1</td>
</tr>
<tr>
<td>Multiple gestation</td>
<td>950 (22)</td>
<td>122 (24)</td>
<td>.2</td>
</tr>
</tbody>
</table>

* Unavailable data: health insurance: NICU, 323 (7%), level I and II nurseries, 81 (16%); education: NICU, 271 (6%), level I and II nurseries, 65 (13%).

---

PEDIATRICS Volume 119, Number 5, May 2007 e1081
significant differences in the incidence of RDS, IVH (1–4 or PVL), early sepsis, NEC (stage II or higher), and BPD between the 2 groups. However, compared with the NICU group, the incidence of late-onset sepsis and ROP was significantly lower in the level I and II nurseries group (P < .0001).

Table 3 describes the neonatal hospital stay characteristics for NICU and level I and II nurseries groups. The percentage of outborn infants was significantly higher in the level I and II nurseries group (26%) compared with the NICU group (8%; P < .0001). Within the level I and II nurseries group, infant weight at transfer from the NICU to level I and II nurseries was 1.6 ± 0.7 kg (range: 0.6–7.2) and the CA at transfer was 34.7 ± 4.5 weeks (range: 24.9–69.7). At the time of discharge from the hospital, the level I and II nurseries infants were 160 g lighter (P = .003) and had a gestational age 0.3 week lower (P = .02) compared with the NICU infants. Also, infants who were transferred to level I and II nurseries had 5.3 days shorter total hospitalization than those who were discharged from the hospital from the NICU (P < .0001). Length of hospitalization ranged from a low of 19 days for a NICU infant to a high of 670 days for a level I and II nurseries infant. One NICU infant with intrauterine growth restriction was discharged at 19 days and 36.5 weeks.

Table 4 shows the follow-up rates at 18 to 22 months’ CA and postdischarge rehospitalization and death rates in the NICU and level I and II nurseries groups. Of the 4795 surviving infants who were eligible for follow-up, 4198 (88%) completed the follow-up assessment. The follow-up rate was significantly lower in the level I and II nurseries group compared with the NICU group (78% vs 89%; P < .0001). Infants in the NICU group were more likely to be seen within the study window of 18 to 22 months’ CA (P < .0004). The incidence of rehospitalization before the follow-up visit was significantly higher in the level I and II nurseries group (56%) compared with the NICU group (49%; P = .02); however, rehospitalization in the first year of life after discharge home was similar between the 2 groups (P = .8). Also, there was a higher death rate between discharge and the follow-up visit at 18 to 22 months’ CA for infants who were discharged from level I and II nurseries compared with those who were discharged from the NICU (4% vs 2%; P = .04). The death rate, however, in the first 6 months after discharge home was similar in the 2 groups (1% vs 2.3%; P = .2).

Table 5 shows the causes of rehospitalization and death after discharge for the NICU and level I and II nurseries groups. A significantly higher percentage of infants in level I and II nurseries group were rehospital-
ized for surgical issues compared with the NICU group ($P = .04$). Also, 52 (14%) infants in the level I and II nurseries group compared with 395 (10%) in NICU group received hernia repair surgery after discharge home ($P = .04$; data not shown). Among infants who did not have hernia repair surgery, the rehospitalization rate in the 2 groups was similar (level I and II nurseries: 161 [50%]; NICU: 1559 [46%]; $P = .16$). Although infants in the level I and II nurseries and NICU groups had similar rates of discharge home on oxygen therapy (134 [30%] and 1212 [28%], respectively), the rate of rehospitalization among infants who were discharged from the hospital on oxygen was significantly higher in the level I and II nurseries group compared with the NICU group (74 [71%] vs 621 [59%]; $P = .02$); no difference was noted in rehospitalization rates for infants who were discharged from the hospital on room air (level I and II nurseries: 115 [49%]; NICU: 1241 [45%]; $P = .4$). Respiratory syncytial virus prophylaxis was reported for 192 (51%) infants in the level I and II nurseries group and for 2137 (56%) infants in NICU group ($P = .07$).

There were a total of 97 postdischarge deaths. Among the infants who died in the 2 study groups, the only differences between NICU and level I and II nurseries groups were birth weight (780 ± 144 g vs 695 ± 141 g; $P = .04$), outborn status (8 [10%] vs 5 [33%]; $P = .03$), and length of NICU stay (126 ± 71 vs 76 ± 41; $P = .003$), respectively. In the level I and II nurseries group, 4 infants died at the levels I and II hospital before discharge home: 2 deaths were secondary to NEC within 2

### Table 4: Follow-up Rates and Postdischarge Rehospitalization and Death Rates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>NICU, n (%)</th>
<th>Level I and II Nurseries, n (%)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>4392</td>
<td>504</td>
<td></td>
</tr>
<tr>
<td>Death after discharge*</td>
<td>82 (2)</td>
<td>15 (4)</td>
<td>.04</td>
</tr>
<tr>
<td>Death rate in first 6 mo after discharge home*</td>
<td>51 (1)</td>
<td>9 (2.3)</td>
<td>.2</td>
</tr>
<tr>
<td>Survivors eligible for follow-up assessment†</td>
<td>4310 (98)</td>
<td>485 (96)</td>
<td></td>
</tr>
<tr>
<td>Lost to follow-up</td>
<td>489 (11)</td>
<td>108 (22)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Follow-up assessment‡</td>
<td>3821 (89)</td>
<td>377 (78)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Follow-up visit in 18- to 22-mo window‡</td>
<td>3362 (88)</td>
<td>306 (82)</td>
<td>&lt;.0004</td>
</tr>
<tr>
<td>Rehospitalization before follow-up visit‡</td>
<td>1873 (49)</td>
<td>209 (56)</td>
<td>.02</td>
</tr>
<tr>
<td>Rehospitalization in first year of life‡</td>
<td>1358 (37)</td>
<td>136 (37)</td>
<td>.8</td>
</tr>
</tbody>
</table>

*Denominator for these variables is the number of infants who received follow-up assessment.
†Infants who were eligible for follow-up included infants who were discharged home minus infants who died after discharge.
‡Denominator for this variable is infants who were eligible for follow-up.
§Denominator for these variables is the number of infants who received follow-up assessment and data on timing of rehospitalization was available.

### Table 5: Causes for Rehospitalizations and Death After Discharge Home

<table>
<thead>
<tr>
<th>Parameter</th>
<th>NICU, n (%)</th>
<th>Level I and II Nurseries, n (%)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of rehospitalizations*</td>
<td>3809</td>
<td>375</td>
<td></td>
</tr>
<tr>
<td>Causes of rehospitalization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respiratory</td>
<td>1008 (26)</td>
<td>106 (28)</td>
<td>.5</td>
</tr>
<tr>
<td>Central nervous system</td>
<td>83 (2)</td>
<td>9 (2)</td>
<td>.9</td>
</tr>
<tr>
<td>Surgical</td>
<td>580 (15)</td>
<td>73 (19)</td>
<td>.04</td>
</tr>
<tr>
<td>Infection</td>
<td>426 (11)</td>
<td>32 (9)</td>
<td>.1</td>
</tr>
<tr>
<td>Growth and nutrition</td>
<td>149 (4)</td>
<td>13 (3)</td>
<td>.8</td>
</tr>
<tr>
<td>Environmental</td>
<td>10 (0.3)</td>
<td>1 (0.2)</td>
<td>.999</td>
</tr>
<tr>
<td>Other</td>
<td>123 (3)</td>
<td>23 (6)</td>
<td>.005</td>
</tr>
<tr>
<td>No. of deaths</td>
<td>82</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Causes of death</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congenital malformation</td>
<td>1 (1)</td>
<td>1 (7)</td>
<td>.3</td>
</tr>
<tr>
<td>BPD</td>
<td>12 (15)</td>
<td>2 (13)</td>
<td>.999</td>
</tr>
<tr>
<td>Infection</td>
<td>5 (6)</td>
<td>1 (7)</td>
<td>.999</td>
</tr>
<tr>
<td>Trauma</td>
<td>2 (2)</td>
<td>0</td>
<td>.999</td>
</tr>
<tr>
<td>SIDS</td>
<td>13 (16)</td>
<td>6 (40)</td>
<td>.08</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>3 (4)</td>
<td>0</td>
<td>.999</td>
</tr>
<tr>
<td>Other</td>
<td>25 (31)</td>
<td>3 (20)</td>
<td>.4</td>
</tr>
<tr>
<td>Child abuse</td>
<td>1 (1)</td>
<td>0</td>
<td>.999</td>
</tr>
<tr>
<td>Unknown</td>
<td>20 (24)</td>
<td>2 (13)</td>
<td>.5</td>
</tr>
</tbody>
</table>

*The total number of hospitalization for all infants who were readmitted for any of the causes.
weeks of transfer, 1 was related to BPD at ~2 months after transfer, and 1 infant’s cause of death was unknown. Infants in the NICU group died at a mean age of 287 days (range: 62–810) and infants in the level I and II nurseries group died at a mean age of 267 days (range: 146–545; \( P = .81 \)). The causes of death included BPD, infection, sudden infant death syndrome (SIDS), congenital malformation, and congestive heart failure. (Table 5) Also, there was a trend for increased deaths secondary to SIDS in the level I and II nurseries group (6 [40%]) compared with the NICU group (13 [16%]; \( P = .08 \)).

Neurodevelopmental and growth outcomes at the follow-up visit 18 to 22 months’ CA are shown in Table 6. The mean Bayley MDI score in the level I and II nurseries group was 2.3 points higher than in the NICU group (\( P = .01 \)). NDI was identified in 38% of infants who were cared for in the NICU and 32% of those who were transferred to level I and II nurseries (\( P = .03 \)). Growth parameters were similar among NICU and level I and II nurseries groups at follow-up.

Table 7 shows the multiple logistic regression estimates of the impact of transfer to level I and II nurseries and outborn status on NDI, MDI and PDI, and rehospitalization between discharge and the follow-up visit at 18 to 22 months’ CA. After adjustment for standard covariates, transfer to level I and II nurseries was not associated with an increased risk for NDI or for MDI or PDI <70. Although transfer to a level I or II nursery was associated with increased risk for postdischarge rehospitalization between discharge and the follow-up visit at 18 to 22 months’ CA (odds ratio [OR]: 1.5; 95% confidence interval [CI]: 1.1–1.9), it was not associated with an increased risk for rehospitalization in the first year of life (OR: 1.1; 95% CI: 0.8–1.5). Also, after adjustment for the standard confounders, transfer to level I and II nurseries did not increase the risk for postdischarge death in the first 6 months of life (OR: 2.0; 95% CI: 0.7–5.9) or before the follow-up visit at 18 to 22 months’ CA (OR: 2.1; 95% CI: 0.8–5.1; regression not shown). Outborn status, however, increased the likelihood of NDI (OR: 1.4; 95% CI: 1.1–1.8) and MDI <70 (OR: 1.4; 95% CI: 1.1–1.8) and rehospitalization between discharge and 18 to 22 months’ CA (OR: 1.6; 95% CI: 1.2–2.0).

Analyses were completed to assess the characteristics of infants who were followed with those who were lost to follow-up at 18 to 22 months’ CA. Of 4795 surviving eligible infants, 597 were lost to follow-up. Infants who were lost to follow-up were more likely to be outborn (14% vs 10%; \( P = .002 \)), born to mothers who were younger than 20 years (19% vs 15%; \( P = .0002 \), and publicly insured (69% vs 64%; \( P = .03 \)). Also, infants who were lost to follow-up were of higher gestational age at birth (26.5 ± 2.1 vs 26.2 ± 2.0; \( P = .009 \)) and had shorter NICU (88 ± 46 vs 94 ± 44; \( P < .0001 \)) and hospital stay (94 ± 43 vs 97 ± 44; \( P = .02 \)) than those who were followed up. There were no significant differences in neonatal morbidities between the 2 groups except higher incidence of ROP in the group that was followed-up (72%) compared with the group that was lost to follow-up (67%; \( P = .03 \)). After adjustment for standard covariates, only higher gestational age (OR: 1.1; 95% CI: 1.0–1.2) was associated with increase risk for loss to follow-up at 18 to 22 months’ CA. Transfer to level I and II nurseries did not independently contribute to loss to follow-up.

### Table 6: Neurodevelopmental and Growth Outcomes at the Follow-up Visit 18 to 22 Months’ CA

<table>
<thead>
<tr>
<th>Parameter</th>
<th>NICU</th>
<th>Level I and II Nurseries</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA at evaluation, mean ± SD (range), mo</td>
<td>19.9 ± 2.1 (14.2–34.8)</td>
<td>20.4 ± 2.6 (13.5–36.5)</td>
<td>.0002</td>
</tr>
<tr>
<td>Neurodevelopment assessment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDI, mean ± SD</td>
<td>78.7 ± 18.1</td>
<td>81.0 ± 18.3</td>
<td>.01</td>
</tr>
<tr>
<td>PDI, mean ± SD</td>
<td>83.0 ± 18.3</td>
<td>84.9 ± 18.7</td>
<td>.08</td>
</tr>
<tr>
<td>MDD &lt;70, n (%)a</td>
<td>1107 (31)</td>
<td>97 (28)</td>
<td>.1</td>
</tr>
<tr>
<td>PDI &lt;70, n (%)a</td>
<td>738 (21)</td>
<td>65 (19)</td>
<td>.3</td>
</tr>
<tr>
<td>Blindness, n (%)</td>
<td>33 (1)</td>
<td>2 (1)</td>
<td>.8</td>
</tr>
<tr>
<td>Hearing impaired, n (%)</td>
<td>58 (2)</td>
<td>9 (2)</td>
<td>.3</td>
</tr>
<tr>
<td>Cerebral palsy, n (%)</td>
<td>514 (14)</td>
<td>50 (13)</td>
<td>.999</td>
</tr>
<tr>
<td>NDI, n (%)a</td>
<td>1343 (38)</td>
<td>114 (32)</td>
<td>.03</td>
</tr>
<tr>
<td>Growth assessment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight, mean ± SD, kg</td>
<td>10.4 ± 2.2</td>
<td>10.4 ± 1.5</td>
<td>.5</td>
</tr>
<tr>
<td>Weight &lt;10%, n (%)</td>
<td>1818 (48)</td>
<td>180 (48)</td>
<td>.9</td>
</tr>
<tr>
<td>Length, mean ± SD, cm</td>
<td>80.6 ± 4.5</td>
<td>80.5 ± 4.6</td>
<td>.9</td>
</tr>
<tr>
<td>Length &lt;10%, n (%)</td>
<td>1092 (29)</td>
<td>124 (33)</td>
<td>.06</td>
</tr>
<tr>
<td>Head circumference, mean ± SD, cm</td>
<td>46.7 ± 1.9</td>
<td>46.8 ± 2.0</td>
<td>.4</td>
</tr>
<tr>
<td>Head circumference &lt;10%, n (%)</td>
<td>958 (25)</td>
<td>94 (25)</td>
<td>.9</td>
</tr>
</tbody>
</table>

a Unavailable data: MDI <70: NICU, 286 (7%), level I and II nurseries, 24 (6%); PDI <70: NICU, 318 (8%), level I and II nurseries, 28 (7%); NDI: NICU, 305 (8%), level I and II nurseries, 25 (7%).
DISCUSSION

ELBW infants who were transferred to level I and II nurseries from tertiary care NICUs had similar neurodevelopmental outcomes and growth compared with infants who received continuing care in the NICU until discharge home. Unique demographic characteristics of infants who were cared for in level I and II nurseries were identified. Their mothers are more likely to be white and have private health insurance. The infants who were cared for in level I and II nurseries were more likely to be outborn, spent >1 month in the level I and II nurseries, and were discharged earlier than the NICU infants.

Transfer of clinically stable ELBW infants to level I and II nurseries has potential advantages. In our cohort, infants who were transferred to a level I and II nurseries had a 21-g higher birth weight, 0.4-week higher gestational age, and lower rates of late-onset sepsis and ROP than the NICU group. This suggests that infants who were transferred to level I and II nurseries had a more benign neonatal course, making them candidates for both transfer and earlier discharge to home. In fact, the infants in level I and II nurseries group were discharged 5 days earlier and 160 g lighter in weight than infants in the NICU group. These findings were similar to those that were observed by Pittard et al, who reported that outborn infants who were back-transferred were discharged from the hospital 42 days earlier and with a 353-g lower discharge weight than those who were discharged from the hospital 5 days earlier and 160 g lighter in weight than infants who were discharged from a tertiary center.

Our data suggest that there is a potential for adverse outcomes that are associated with transfer of ELBW infants to level I and II nurseries. In general, ELBW infants remain at risk for late-onset sepsis, respiratory compromise, and failure to thrive during the convalescent period whether it is in the NICU or level I or II nursery. Four infants died after transfer to a level I or II nursery. Although in this study the NICU group rehospitalization rate (49%) and death rate (2%) were similar to those that were reported in previous NICHD network publications, a higher rate of rehospitalization (56%) was reported in infants who were discharged from level I and II nurseries. After adjustment for standard covariates, transfer to level I and II nurseries was associated with an increased risk of rehospitalization between discharge and the follow-up visit at 18 to 22 months’ CA but not in the first year of life. A higher percentage of infants in the level I and II nurseries group were hospitalized for surgical issues, and a significantly higher number of infants who were discharged from level I and II nurseries received hernia surgery after discharge. We speculate that community hospitals where these level I or level II nurseries are located may lack pediatric specialties, including pediatric surgery, requiring that infants be rehospitalized for surgical procedures after discharge home, whereas, NICU infants at tertiary care centers are more likely to have surgical procedures before discharge. Rehospitalization rate was significantly higher for infants who were discharged from the hospital on oxygen from level I and II nurseries compared with those who were discharged from the hospital on oxygen from the NICU. Rehospitalization rates were similar for infants who were discharged from the hospital on room air in the 2 groups. We speculate that infants who were discharged from level I and II nurseries may have decreased access to subspecialty and neonatal follow-up care after discharge. This lack of access would have greater implications for infants with complex medical problems or those who require additional surgery.

After adjustment for standard covariates, the postdischarge death rate was not associated with transfer to level I and II nurseries. The causes of death were similar in the 2 groups, except for the trend for higher incidence of SIDS (40%) in the level I and II nurseries group. Lower birth weight is known to be associated with SIDS.

Although a general framework for different levels of neonatal care exists, there are no standard definitions or classifications for the complexity of care provided at different levels of neonatal units. Some states have policies that designate different levels of care that can be provided by a nursery. However, the interpretation of these designations and the applications of these NICU-related regulations vary considerably. The level of care that is provided by a level I nursery and level III NICUs is fairly consistent. Greater diversity in the level of care is seen in level II nurseries. In our study, 80% of level II nurseries and only 20% of level I nurseries that received transfers from the network NICU centers were staffed by a neonatologist, and 50% of level II nurseries provided ventilatory care in the unit. Also, none of the level I or level II nurseries had a lower limit for infant weight or gestational age for admission. In this cohort, at least 1 infant had a transfer weight of 600 g and another was transferred at a CA of 25 weeks (confirmed by centers). This suggests that there is a degree of variability in the level of intensive care that is provided at different community hospital nurseries that are affiliated with the

| TABLE 7 | Adjusted Risk Factors for NDI and for MDI and PDI <70 at Follow-up Visit and Rehospitalization Before Follow-up Visit |
|-----------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|
| Risk Factor      | NDI, OR (95% CI) | MDI < 70, OR (95% CI) | PDI < 70, OR (95% CI) | Rehospitalization, OR (95% CI) |
| Transfer to level I and II nurseries | 0.84 (0.62–1.14) | 0.95 (0.69–1.31) | 0.98 (0.68–1.40) | 1.46 (1.11–1.91) |
| Outborn          | 1.37 (1.06–1.77) | 1.37 (1.05–1.78) | 1.27 (0.95–1.70) | 1.56 (1.23–1.98) |

Factors included neonatal centers, outborn, transfer to level I and II nurseries, multiple births, maternal education, insurance type, gender, birth weight, gestational age, and neonatal morbidities.
network centers. We speculate that there may be differences in the medical management and comprehensive discharge planning that are provided to infants who are discharged from different level I and level II nurseries that may affect hospital readmission rates. The characteristics of the level of care that is provided to ELBW infants as well as the discharge planning and postdischarge management at community hospitals requires additional evaluation.

Infants in level I and II nurseries had higher birth weight and gestational age and lower neonatal morbidities. NDI is known to be associated with lower birth weight and neonatal morbidities. In our study, after adjustment for demographic and neonatal characteristics, NDI was not affected by transfer to level I and II nurseries; delivery outside the tertiary center was associated with increased risk for NDI and for MDI <70. Higher mortality and morbidity in infants who were born outside a tertiary center have been reported; however, an association with worse NDI and MDI was not reported previously.

Transfer of infants to levels I and II hospitals within the community may decrease parental stress by facilitating better parental and sibling bonding and allowing the continuing care physician to be involved in the medical issues that are specific to the infant. However, transfer also may lead to increased stress on the parent. Family reaction to transfer may be either positive or negative. This depends on many factors, including the transfer experience of staff, the quality of communication between the staff at the 2 hospitals, and the availability of support services for the parents. Lack of communication between the 2 hospitals may result in inadequate referrals and follow-up for needed services such as early intervention, developmental assessment, and subspecialty clinics. In this study, 78% of infants who were cared for in level I and II nurseries compared with 89% of NICU infants were assessed at follow-up at 18 to 22 months' CA. Similar to other studies, in this cohort, outborn infants who were born to younger mothers with public insurance, had higher gestational age, and had shorter hospitalization were more likely to be noncompliant for follow-up. Infants in the level I and II nurseries group also were more likely to receive follow-up assessment beyond the 18- to 22-months CA window than those in the NICU group. Increased geographic distance and limited access to specialty and support services may be some of the additional factors that are difficult to measure but may have contributed toward the loss to follow-up.

CONCLUSIONS

In this ELBW cohort, infants who were transferred to level I and II nurseries had fewer neonatal morbidities, were discharged earlier, and had similar NDI and growth compared with the NICU only infants. However, infants who were cared for in level I and II nurseries were more likely to be rehospitalized and were at an increased risk for loss to follow-up. Establishment of consistent guidelines for neonatal management and comprehensive discharge planning by community hospital nurseries in partnership with the tertiary center NICUs may improve follow-up compliance and reduce rehospitalizations.

ACKNOWLEDGMENTS

Members of the NICHD Neonatal Research Network were as follows: follow-up center, principal investigator, follow-up principal investigator, network coordinator, follow-up coordinator (grants): Brown University, William Oh, MD, Betty Vohr, MD, Angelita Hensman, RNC, Lucy Noel, RNC (U10 HD27904); Case Western Reserve University, Avroy A. Fanaroff, MB, BCh, Dee Wilson, MD, Nancy Newman, RN, Bonnie Singer, RN (U10 HD21364); Duke University, Ron Goldberg, MD, Ricky Goldstein, MD, Melody Lohmeyer, Melody Lohmeyer (U10 HD40492); Emory University, Barbara J. Stoll, MD, Barbara J. Stoll, MD, Ellen Hale, RNC, BS, Ellen Hale, RNC, BS (U10 HD27851, M01 RR0039); Harvard University, Ann R. Stark, MD, Ann R. Stark, MD, Kerri Fournier, RN (U10 HD34167, M01 RR02635, M01 RR02172, M01 RR01032); Indiana University, James A. Lemons, MD, Anna Dusick, MD, DeeDee Appel, RN, Leslie Richards, RN (U10 HD27856, M01 RR00750); Stanford University, David K. Stevenson, MD, Susan Hintz, MD, Bethany Ball, RN, Bethany Ball, RN (U10 HD27880, M01 RR00070); University of Alabama, Waldemar A. Carlo, MD, Myriam Peralta, MD, Monica Collins, RN, Vivien Phillips (U10 HD34216); University of Cincinnati, Edward F. Donovan, MD, Jean Steichen, MD, Cathy Grisby, RN, Teresa Gratton, PA (U10 HD27853, M01 RR08084); University of Miami, Shahnaz Durara, MD, Charles Bauer, MD, Ruth Everett, RN, Mary Allison, RN (U10 HD21397); University of New Mexico, Lu-Ann Papile, MD, Lu-Ann Papile, MD, Conra Backstrom, RN (U10 HD27881, M01 RR00997); University of Tennessee, Sheldon B. Korones, MD, Kimberly Yolton, PhD, Tina Hudson, RN (U10 HD21415); University of Texas–Dallas, Abbot R. Laptook, MD, Roy Heyne, MD, Susie Madison, RN, Jackie Hickman, RN (U10 HD40689); University of Texas–Houston, John E. Tyson, Brenda Morris, Georgia McDavid, Georgia McDavid (U10 HD21373); University of California–San Diego, Neil Fiener, MD, Yvonne Vaucher, MD, Wade Rich, Martha Fuller (U10 HD40461); Wayne State University, Seetha Shankaran, MD, Yvette Johnson, MD, Gerry Muran, BSN, Debbie Kennedy, RN (U10 HD21385); Wake Forest University, Michael O’Shea, MD, Robert Dillard, MD, Nancy Peters, Barbara Jackson (U10 HD40498); Yale University, Richard A. Ehrenkranz, MD, Richard A. Ehrenkranz, MD, Pat Gettner, RN, Elaine Romano, MSN (U10 HD27871, M01 RR06022); Rochester University, Dale L. Phelps, MD, Gary Myers, MD, Linda Reubens,
REFERENCES

ARTICLE

Childhood Cancer and Birthmarks in the Collaborative Perinatal Project

Kimberly J. Johnson, MPH\textsuperscript{a}, Logan G. Spector, PhD\textsuperscript{a}, Mark A. Klebanoff, MD, MPH\textsuperscript{b}, Julie A. Ross, PhD\textsuperscript{a}

\textsuperscript{a}Department of Pediatrics, University of Minnesota, Minneapolis, Minnesota; \textsuperscript{b}National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland

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

ABSTRACT

OBJECTIVE. Three previous retrospective studies noted a positive association between birthmarks and childhood cancer. The objective of this study was to determine whether the incidence of cancer is increased in children with birthmarks relative to those without birthmarks using data from the Collaborative Perinatal Project cohort, a large, prospective study.

METHODS. Our study population comprised 49,503 US children who were born between 1959 and 1968. Birthmarks were documented as definite or suspected during the first year through history or medical examinations and included hemangiomas (port-wine, strawberry, or cavernous), pigmented nevi, lymphangiomas, and café-au-lait spots. The association between birthmarks and childhood cancer was determined using Cox proportional hazards regression.

RESULTS. In the Collaborative Perinatal Project, 2,505 individuals had a documented definite or suspected birthmark, including 7 of 47 children who developed cancer. Birthmarks were associated with a significant increase in the risk for cancer. There was a slight attenuation of the risk estimate when cases that were diagnosed in the first year of life were excluded. No specific childhood malignancies were notably affected by birthmarks.

CONCLUSIONS. Although this study was based on a small number of cases, we found birthmarks to be in excess in children who received a diagnosis of cancer using prospective data. These findings provide additional support for the possibility of a shared etiology between birthmarks and childhood cancer that could offer insight into the pathogenesis of pediatric malignancy.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2315
doi:10.1542/peds.2006-2315

Key Words

birthmarks, childhood cancer, etiology, risk factors

Abbreviations

BWS—Beckwith-Wiedemann syndrome
NF1—neurofibromatosis type I
ALL—acute lymphoblastic leukemia
CPP—Collaborative Perinatal Project
HR—hazard ratio
CI—confidence interval
SES—socioeconomic status
AML—acute myelogenous leukemia
IGF—insulin-like growth factor

Accepted for publication Nov 8, 2006
Address correspondence to Julie A. Ross, PhD, Department of Pediatrics, University of Minnesota, 420 Delaware St SE, MMC 422, Minneapolis, MN 55455. E-mail: ross@epi.umn.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
It is widely recognized that genetic syndromes and chromosomal anomalies such as Down syndrome, Beckwith-Wiedemann syndrome (BWS), and neurofibromatosis type I (NF1) are associated with a markedly increased risk for childhood malignancies. Studies also have suggested that children with minor malformations or variants, including “birthmarks,” may have an increased risk, and these could be a marker of “altered prenatal development.” A significant increased frequency of birthmarks in children with acute lymphoblastic leukemia (ALL) and rhabdomyosarcoma has been reported in 3 studies. To provide additional data on the association between childhood cancer and the presence of skin lesions that are detected during infancy, we explored whether birthmarks were associated with childhood cancer in the Collaborative Perinatal Project (CPP), a large, prospective cohort study.

METHODS

Study Population

The CPP cohort was assembled between 1959 and 1966 with the main objective of investigating the causes of neurologic defects. Complete details on the study have been described elsewhere. Briefly, ~45,000 women were enrolled during their pregnancies at 12 academic institutions throughout the United States. A total of 59,843 pregnancies that resulted in 54,795 live births occurred. Data were collected at numerous time points by trained interviewers and examiners using standardized methods. During the prenatal phase, extensive information was collected on the mother’s health history, pregnancy-related events, social and demographic factors, and family medical histories. During and after delivery, trained observers were present to collect standard information and to document any significant events. Children were examined several times during the newborn period and at 4, 8, and 12 months during the first year. For most individuals in the cohort, a summary form was filled out to record all events, illnesses, and conditions that occurred between discharge from the birth hospital and the first birthday. Investigators continued to collect data at several time points through the age of 7 or 8 years (depending on the institution).

Our study population for the analysis of the association between birthmarks and childhood cancers was composed of 49,503 children who were born alive and had documented information on birthmarks that were observed during the first year of life. Fifty-one cancers that were identified during 8 years of follow-up were described previously by Klebanoff et al. Four of the 51 children who received a diagnosis of cancer had no documentation on the existence of birthmarks and were excluded from this analysis. Children were examined for the presence or absence of infantile hemangiomas (strawberry and cavernous), port-wine stains, pigmented nevi, lymphangiomas, and café-au-lait spots at examinations that occurred during the first year of life. Examination forms allowed the examiners to note birthmarks that they considered to be present beyond doubt (definite birthmarks) and those for which they had some degree of doubt (suspected birthmarks). Examiners were instructed to exclude nevus simplex (also known as stork bites or salmon patch), the most common skin lesion that is observed during infancy, and not to record pigmented nevi or café-au-lait spots that were <3 cm or fewer than 6.

Statistical Analysis

All statistical analyses were performed using SAS 9.1 (SAS Institute, Cary, NC). The Kaplan-Meier method was used to estimate the cumulative hazard of childhood cancer with respect to the presence of birthmarks. We used Cox proportional hazards regression modeling to calculate hazard ratios (HRs), 95% confidence intervals (CIs), and likelihood ratio $\chi^2$ statistics. Person-time was calculated as the number of months from birth to cancer diagnosis or the last date of follow-up as documented in electronic records, whichever came first. The maximum follow-up was 96 months. Because birthmarks were noted throughout the first year of life and not necessarily at birth, analyses were repeated in which person-time started at age 1 year (excluding those who had person-times of <12 months) to disallow the possibility that birthmarks were discovered coincident with cancer diagnosis.

We considered whether the following potential confounders that have previously been linked to childhood cancer were associated with cancers that were diagnosed in the CPP in Cox proportional hazards regression models: socioeconomic status (SES; 3 categories), birth weight ($\leq2500$ g, $2501$–$4000$ g, or $\geq4000$ g), maternal race (white, black, or other), gender, maternal education (less than high school graduate, high school graduate, more than high school graduate), maternal age ($\geq35$ years), and gestational age (<37 weeks). For analyses that examined the association between childhood cancer and birthmarks, we created binary variables to describe the presence or absence of 1 or more of these characteristics using 2 birthmark definitions, 1 that included only definite birthmarks as present and 1 that also included suspected birthmarks. We used the change in estimate method using a cutoff of 10% to evaluate whether the association between childhood cancer and birthmarks was confounded by each of the variables described.

RESULTS

A total of 47 and 38 incident cancer cases were diagnosed between 0 and 8 years and 1 and 8 years, respectively, in children who had documentation about birthmarks (Table 1). The 3 most common cancers were leukemia, central nervous system tumors, and Wilms’
tumor, as is the normal pattern of childhood cancers in this age group.19

No significant associations (as determined by the likelihood ratio $\chi^2$ test comparing models with and without the covariate) were observed for any of the following risk factors that are known to be associated with some childhood cancers17 (Table 2): SES, birth weight, maternal race, gender, maternal education, or maternal or gestational age. However, a general tendency was apparent for the HR of childhood cancer to decrease with lower SES and maternal education and for an increased hazard of childhood cancer associated with maternal white ethnicity.

Approximately 15% of cases ($n = 7$) and 5% of noncases ($n = 2498$) had a definite or suspected birthmark documented during the first year of life (Table 3), with strawberry hemangiomas and port-wine stains being the most common in both groups. A similar percentage of birthmarks in cases (2 [29%] of 7) and noncases (665 [24%] of 2721) were documented as suspected. The median follow-up time in children with and without documented birthmarks was identical at 96 months.

Figure 1 shows the cumulative hazard of childhood cancer over time with respect to birthmark group. The log-rank test of equality over birthmark strata showed that children with birthmarks were significantly more likely to have a subsequent diagnosis of cancer than those without birthmarks during the follow-up period ($P = .0028$).

Table 4 summarizes the analysis of the association between childhood cancer and birthmarks. Having a documented definite or suspected birthmark was associated with a significantly increased hazard of cancer (HR: 3.19; 95% CI: 1.43–7.12) when all cases ($n = 47$) were included. When individuals with suspected birthmarks were excluded ($n = 485$), the HR decreased to 2.81 but was still significant (95% CI: 1.11–7.13). The HR for the association between childhood cancer and birthmarks when individuals with follow-up times of $<1$ year were excluded and both definite and suspected birthmarks were included did not materially change the results (HR: 2.74; 95% CI: 1.07–7.02). Excluding individuals with follow-up times of $<1$ year resulted in an HR of 2.0 that was no longer significant (95% CI: 0.62–6.62) for definite birthmarks only. Adjustment for potential founders (SES, birth weight, maternal race, gender, maternal education, maternal age, and gestational age) did not materially change the results.

Table 5 shows the types of birthmarks (definite and

### Table 1: Cancers Diagnosed in the CPP Cohort ($N = 49503$)

<table>
<thead>
<tr>
<th>Type of Cancer</th>
<th>No. Diagnosed Between 0 and 8 y</th>
<th>No. Diagnosed Between 1 and 8 y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leukemia</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>CNS</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Wilms’ tumor</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Neuroblastoma</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Retinoblastoma</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>RMS</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hepatoblastoma</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Fibrosarcoma</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>38</td>
</tr>
</tbody>
</table>

CNS indicates central nervous system; RMS, rhabdomyosarcoma.

### Table 2: HRs of Childhood Cancer by Potential Confounding Factors in the CPP ($N = 49503$)

<table>
<thead>
<tr>
<th>Covariate</th>
<th>No. of Cases</th>
<th>Child-mo HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (low)</td>
<td>14</td>
<td>1.565741</td>
<td>0.58</td>
</tr>
<tr>
<td>2 (middle)</td>
<td>15</td>
<td>1.232521</td>
<td>0.78</td>
</tr>
<tr>
<td>3 (high)</td>
<td>18</td>
<td>1.154785</td>
<td>1.0</td>
</tr>
<tr>
<td>Birth weight, g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\leq2500$</td>
<td>3</td>
<td>420960</td>
<td>0.58</td>
</tr>
<tr>
<td>2501–4000</td>
<td>42</td>
<td>3418797</td>
<td>1.0</td>
</tr>
<tr>
<td>$&gt;$4000</td>
<td>2</td>
<td>207310</td>
<td>0.78</td>
</tr>
<tr>
<td>Maternal race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>25</td>
<td>1823992</td>
<td>1.0</td>
</tr>
<tr>
<td>Black</td>
<td>21</td>
<td>2014894</td>
<td>0.77</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>213038</td>
<td>0.32</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>24</td>
<td>2046173</td>
<td>1.0</td>
</tr>
<tr>
<td>Female</td>
<td>23</td>
<td>2005598</td>
<td>0.98</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>24</td>
<td>2288131</td>
<td>1.0</td>
</tr>
<tr>
<td>High school</td>
<td>16</td>
<td>1249049</td>
<td>1.22</td>
</tr>
<tr>
<td>More than high school</td>
<td>7</td>
<td>452503</td>
<td>1.47</td>
</tr>
<tr>
<td>Maternal age, y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$&lt;35$</td>
<td>44</td>
<td>3798441</td>
<td>1.0</td>
</tr>
<tr>
<td>$\geq35$</td>
<td>3</td>
<td>253438</td>
<td>1.03</td>
</tr>
<tr>
<td>Gestational age, wk</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$&lt;37$</td>
<td>39</td>
<td>3495459</td>
<td>1.0</td>
</tr>
<tr>
<td>$\geq37$</td>
<td>8</td>
<td>556420</td>
<td>1.29</td>
</tr>
</tbody>
</table>

a A total of 1201 observations had missing data on SES.
b A total of 64 observations had missing birth weights.
c Two observations had missing data on gender.
d A total of 785 observations had missing data on maternal education.

### Table 3: Frequency of Birthmarks by Type Reported in the CPP ($N = 49503$)

<table>
<thead>
<tr>
<th>Type of Birthmark</th>
<th>Cases ($N = 47$), n (%)</th>
<th>Noncases ($N = 49456$), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Definite</td>
<td>Suspect</td>
</tr>
<tr>
<td>Port-wine stain</td>
<td>2 (4.3)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Strawberry hemangioma</td>
<td>2 (4.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Cavernous hemangioma</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Lymphangiomas</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Café-au-lait spots</td>
<td>0 (0)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Pigmented nevi</td>
<td>1 (2.1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Total</td>
<td>5 (10.6)</td>
<td>2 (4.3)</td>
</tr>
<tr>
<td>Total (definite and suspected)</td>
<td>7 (14.9)</td>
<td>2498 (5.1)</td>
</tr>
</tbody>
</table>

a The total differs from the sum of individual birthmarks because some individuals had >1 type of birthmark.
suspected) that were reported in individuals who received a diagnosis of cancer during the follow-up period. Three cases of acute leukemia had documented definite or suspected birthmarks (1 of each of the types: port-wine stain, strawberry hemangioma, and café-au-lait spots). Two patients with Wilms’ tumor were reported to have had a port-wine stain and a strawberry hemangioma, respectively. A port-wine stain also was reported for a retinoblastoma case, and a hairy pigmented nevi was reported for an individual who received a diagnosis of glioma.

No particular type of birthmark occurred more frequently in children with cancer than without \((P = .1208, \text{Fisher’s exact test})\). Because of the small number of cases with birthmarks, we were not able to determine whether any individual type of birthmark was more common in individual malignancies or to examine whether there was an interaction between birthmarks and other variables, such as gender. However, it is interesting to note that of the 7 individuals who had a birthmark documented either as definite or suspected during the first year, 6 were female (data not shown). In addition, no genetic syndromes that are known to be associated with childhood cancer were recorded in individuals with birthmarks during the follow-up time, including neurofibromatosis, which is linked to both leukemia and café-au-lait spots \(^5\) (data not shown). Of note, however, is that an umbilical hernia was reported in the patient who had Wilms’ tumor and also had a strawberry hemangioma (Table 5). Children with BWS

![FIGURE 1](image-url)  
Kaplan-Meier estimates of the cumulative hazard of childhood cancer in those with birthmarks relative to those without.

<table>
<thead>
<tr>
<th>TABLE 4</th>
<th>Cox Proportional Hazards Regression Modeling of the Association Between Birthmarks and Childhood Cancer in the CPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birthmark</td>
<td>No. of Cases</td>
</tr>
<tr>
<td>Cases diagnosed between 0 and 8 y</td>
<td></td>
</tr>
<tr>
<td>Definite or suspect (n = 49 503)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>40 3 840</td>
</tr>
<tr>
<td>Yes</td>
<td>7 211 448</td>
</tr>
<tr>
<td>Definite only*</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>40 3 840</td>
</tr>
<tr>
<td>Yes</td>
<td>5 171 440</td>
</tr>
<tr>
<td>Cases diagnosed between 1 and 8 years</td>
<td></td>
</tr>
<tr>
<td>Definite or suspectb</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>33 2 767 508</td>
</tr>
<tr>
<td>Yes</td>
<td>5 181 411</td>
</tr>
<tr>
<td>Definite onlyc</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>33 2 767 508</td>
</tr>
<tr>
<td>Yes</td>
<td>3 147 223</td>
</tr>
</tbody>
</table>

* A total of 488 individuals with suspected birthmarks were excluded (n = 49 018).

b Nine individuals who were followed for <1 year were excluded (n = 49 494).

c A total of 494 individuals who had suspected birthmarks or who were followed for <1 year were excluded (n = 49 009).

<table>
<thead>
<tr>
<th>TABLE 5</th>
<th>Types of Birthmarks Found in Individuals With Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birthmark</td>
<td>Cancer Type(s)</td>
</tr>
<tr>
<td>Port-wine stain</td>
<td>Wilms’ tumor</td>
</tr>
<tr>
<td>Strawberry hemangioma</td>
<td>Acute leukemia*</td>
</tr>
<tr>
<td>Café-au-lait spots</td>
<td>Retinoblastoma</td>
</tr>
<tr>
<td>Pigmented nevi</td>
<td>Acute leukemia*</td>
</tr>
<tr>
<td></td>
<td>Glioma</td>
</tr>
</tbody>
</table>

* Suspected birthmarks.
have an increased frequency of umbilical hernias and an elevated risk for embryonal tumors.2,21 This child may have had unrecognized BWS, which was first described in 1963 and was not part of data collection.3

**DISCUSSION**

The results of this prospective study support those of previous studies9–11 that reported an association between childhood malignancy and birthmarks. Children who received a diagnosis of cancer in the CPP were ~3 times more likely to have a documented birthmark than those without this diagnosis.

To our knowledge, no other analyses have been conducted to investigate specifically the association between birthmarks and childhood cancer. As mentioned, 3 studies have reported a significantly increased prevalence of birthmarks in children with cancer compared with control subjects. A study conducted by Mertens et al10 examined the association between congenital abnormalities and childhood leukemia in 2117 cases of ALL and 605 cases of acute myelogenous leukemia (AML). Unexpectedly, more mothers of patients with ALL and AML than of control subjects reported their child’s having large or multiple birthmarks (ALL odds ratio: 1.35 [95% CI: 1.35–1.71]; AML odds ratio 1.89 [95% CI: 1.15–3.11]). The authors noted that no patients were reported to have had NF1. The authors did not note which types of birthmarks were ascertained but did mention that they were defined as “any one larger than a quarter; or 6 or more about the size of a dime.” In a study by Roganovic et al,9 a significantly increased frequency of pigmented nevi and café-au-lait spots was found in patients with childhood leukemia compared with control subjects. These authors also note that NF1 was not found in any of their study participants.9 Finally, a study of rhabdomyosarcoma was conducted to look at the association between parental use of marijuana and cocaine and rhabdomyosarcoma. It is interesting that birthmarks were reported to be significantly associated with rhabdomyosarcoma by the investigators, who were exploring potential confounders for inclusion in their primary analysis.11

Other than the known associations between certain types of birthmarks and cancer (eg, congenital pigmented nevi and melanoma22; café-au-lait spots in patients with NF1 and certain types of cancers, including AML3), there has been little research to provide a biological basis for the connection between the two. A small body of evidence suggests 2 possible links. First, insulin-like growth factor (IGF) system members IGF-1 and IGF-2 have been reported to be involved in the pathogenesis of infantile hemangiomas and also in Wilms’ tumor, ALL, and AML.23–26 It also is of interest to note that the IGF system has been implicated in the cause of several adult cancers, including breast, prostate, colon, and lung cancer.27 In addition, high circulating levels of IGF-1 have been proposed to be a possible explanation for the association between high birth weight and childhood and adult cancers.28 It remains to be determined whether certain types of birthmarks and cancer could stem from aberrations in the IGF system. Second, both solid tumors and vascular anomalies (eg, hemangiomas, port-wine stains) are known to have increased expression of angiogenic proteins, including vascular endothelial growth factor and basic fibroblast growth factor.29–31 In addition, limited evidence suggests that angiogenic proteins are involved in the pathogenesis of hematopoietic cancers.32,33 Whether angiogenic factors that are produced by vascular anomalies could act in a paracrine manner to promote tumor growth is unknown.

The major strength of this study was the rigorous and thorough prospective collection of data in contrast to previous studies that used retrospective designs9–11 and collected data through questionnaire-based methods.10,11 This analysis also had several limitations. First, we had low statistical power because of the small number of cases and the low prevalence of birthmarks in the CPP, making our results particularly sensitive to any misclassification of the exposure. However, any misclassification was likely nondifferential in that it did not depend on disease status; in this specialized case in which the exposure is binary, the resulting bias would favor the null.34 Our results were robust to any misclassification of birthmarks because exclusion of individuals with suspected birthmarks did not materially change the results. Second, we were not able to determine the temporal relationship between ascertainment of the birthmark and the diagnosis for cancers that occurred during the first year of life. Analyses that included cases that were diagnosed during the first year of life could have been subject to bias if cases were scrutinized more carefully for other abnormalities than noncases. To determine whether this type of bias occurred, we conducted analyses that excluded individuals with follow-up times of <1 year. Although the HRs decreased in magnitude, they were still indicative of an increased risk. Finally, because of the small number of childhood cancer cases, it is difficult to discount chance completely as an explanation for these results.

**CONCLUSIONS**

We have reported an association between childhood cancer and birthmarks in the CPP. These results could offer new insight into the potential causes of childhood cancer and suggest a possibly fruitful avenue for additional exploration.

**ACKNOWLEDGMENT**

This work was supported by the Children’s Cancer Research Fund.
REFERENCES

**Healing of Hymenal Injuries in Prepubertal and Adolescent Girls: A Descriptive Study**

John McCann, MD, Sheridan Miyamoto, MSN, FNP, Cathy Boyle, MSN, PNP, Kristen Rogers, PhD

Child and Adolescent Research and Evaluation (CAARE) Diagnostic and Treatment Center, University of California, Davis, Children’s Hospital, Sacramento, California

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

**ABSTRACT**

**OBJECTIVE.** The objective of this study was to identify the healing process and outcome of hymenal injuries in prepubertal and adolescent girls.

**METHODS.** This multicenter, retrospective project used photographs to document the healing process and outcome of hymenal trauma that was sustained by 239 prepubertal and pubertal girls whose ages ranged from 4 months to 18 years.

**RESULTS.** The injuries that were sustained by the 113 prepubertal girls consisted of 21 accidental or noninflicted injuries, 73 secondary to abuse, and 19 “unknown cause” injuries. All 126 pubertal adolescents were sexual assault victims. The hymenal injuries healed at various rates and except for the deeper lacerations left no evidence of the previous trauma. Abrasions and “mild” submucosal hemorrhages disappeared within 3 to 4 days, whereas “marked” hemorrhages persisted for 11 to 15 days. Only petechiae and blood blisters proved to be “markers” for determining the approximate age of an injury. Petechiae resolved within 48 hours in the prepubertal girls and 72 hours in the adolescents. A blood blister was detected at 34 days in an adolescent. As lacerations healed, their observed depth became shallower and their configuration smoothed out. Of the girls who sustained “superficial,” “intermediate,” or “deep” lacerations, 15 of 18 prepubertal girls had smooth and continuous appearing hymenal rims, whereas 24 of 41 adolescents’ hymens had a normal, “scalloped” appearance and 30 of 34 had no disruption of continuity on healing. The final “width” of a hymenal rim was dependent on the initial depth of the laceration. No scar tissue formation was observed in either group of girls.

**CONCLUSIONS.** The hymenal injuries healed rapidly and except for the more extensive lacerations left no evidence of a previous injury. There were no significant differences in the healing process and the outcome of the hymenal injuries in the 2 groups of girls.
The evaluation of the female child or adolescent who is suspected of having been sexually abused traditionally focuses on the condition of the hymenal membrane. Examinations that are performed shortly after an assault may disclose findings that are consistent with a recent injury. However, if an assault had taken place sometime in the past, then signs of trauma may have faded as the injuries healed. It is the interpretation of the nonacute examination findings that continues to be debated.

Until recently, there has been relatively little information in the medical literature regarding the healing process and the outcome of a female genital injury, particularly in the case of the prepubertal child. The similarity between naturally occurring variations and a hymenal configuration that results from an injury further complicates the interpretation of a finding.

As a hymenal laceration heals, it may or may not leave evidence of the previous injury. McCann et al observed that hymenal lacerations healed rapidly and "smoothed off" over time. Kerns et al used the term "concave" to describe the multiple variations of healed hymenal lacerations. A recent article by Heppenstall-Heger et al reported that "partial" hymenal tears in 8 preadolescent girls healed "completely," whereas 5 had a "shallow notch" at the site of their injury. The findings that "persisted" were those that were created by a transection. Berenson et al reported that the only child with a hymenal rim 1 mm or less in width had a history of penile penetration. Unfortunately, there were too few cases to determine the statistical significance of that finding. Adams, in a commentary on normal hymenal findings, stated, "If there is a clear rim of hymenal tissue in the posterior aspect of the orifice, and the free edge of the hymen can be followed visually at least from the 9 o'clock to the 3 o'clock positions, when the patient is supine, this is likely to be a normal finding." This project was designed to explore further the findings in these and other reports while determining whether there is a "pattern," a "time sequence," or a "marker" in the healing process that could be used to determine the age of a hymenal injury.

**METHODS**

**Recruitment**
The patients of this multicenter study were recruited from medical centers throughout the United States. The majority of the cases were obtained through the use of the Heller Honorary Society’s List Server. The members of this society are recognized for their expertise in the evaluation and treatment of abused and neglected children and adolescents. The participants were asked to provide pertinent medical information and photographs of any female child or adolescent who had sustained a recent genital injury from any cause. Patients from birth through 17 years of age were eligible. In addition to the photograph documentation of an injury, all patients were required to have at least 1 follow-up examination. Because this was a retrospective, convenience-sample study, the period between an injury and the follow-up examination was not uniform. Each center’s institutional review board authorized its center’s participation in the project.

**Historical Information**
The participants provided the authors with a summary of the portion of a patient’s medical chart that pertained to the genital injury. The information requested included the individual’s birth date and ethnicity, the examiner’s opinion as to the cause of injury, and the examination method used. The time and date of all examinations became part of a computer-generated database. The patient’s computerized medical chart and photographs were assigned a number to protect the individual’s identity.

**Photographic Documentation**
Photographic documentation by the participating institutions was achieved through the use of a variety of recording methods. The most common recording device was a 35-mm camera with either a macro lens or a camera that was mounted on a colposcope. Several centers provided images that had been captured through the use of digital or video cameras. Prints of the images were provided by each center.

**Analysis of the Patients**
The patients were examined by a variety of methods. These included the supine, labial separation method; the supine, labial traction technique; and the prone, knee-to-chest position approach. When a patient had been examined by >1 method, we divided the photographs into separate envelopes on the basis of the method used. Each photograph was evaluated in the presence of all 3 of the medical examiner authors. During the evaluation, the authors were blinded to the history that had been provided by the medical examiners from the contributing centers. An agreement by all 3 medical authors was required before the interpretation of a finding was recorded on a worksheet illustration and entered into the databank.

**Analysis of the Patients**
We divided the patients into 2 groups on the basis of the hormonal effect on the hymen. The first group consisted of the girls whose hymen showed no estrogen effect. Their hymens tended to be thin, delicate membranes with relatively smooth edges. The few girls who were younger than 3 years and retained some visual evidence of endogenous estrogen were placed in the first group, which is referred to as prepubertal girls. The second
group consisted of the older girls whose hymen did show an estrogen effect. Their hymens tended to be thicker and more redundant and frequently had scalloped edges. This second group is referred to as pubertal adolescents.

**Types of Hymenal Injuries**

We subdivided the hymenal injuries into abrasions, contusions, and lacerations. Evidence of a contusion included the presence of blood blisters, edema, hematomas, petechiae, and submucosal hemorrhages. The abrasions and contusions were classified further as to their size and color. We subdivided the hymenal membrane into quadrants for purposes of identifying the location of any abrasions or contusions. Regardless of the examination position used, the location of a hymenal laceration was recorded as though the patient were in a supine position.

The lacerations were categorized according to both depth and configuration. The classification system that was used for the depth of the hymenal lacerations is similar to the one used by Berenson et al.\(^4\) in their report on hymenal injury findings. We determined the depth of a hymenal laceration by comparing the width of the lacerated portion of the hymen with the width of an adjacent, uninjured portion of the membrane.

The depths of the hymenal lacerations were divided into (1) those that penetrated to <50% of the width of the membrane (superficial), (2) those that were approximately halfway through the membrane (intermediate), (3) those that went beyond the midpoint of the membrane (deep), (4) those that extended to the base (attachment) of the hymenal membrane (transection), and (5) those that went through the hymenal membrane attachment into the surrounding tissues (transection with an extension) (see Appendix).

The configuration system that was used to categorize a hymenal laceration’s shape came from the authors’ previous observation that the configuration of a hymenal laceration seemed to change as a laceration healed.\(^2\) It had been noted that acute hymenal lacerations had a sharper V-shaped configuration, whereas the healed lacerations had a smoother, U-shaped appearance. We used this observation as the basis for exploring the possibility that these changes could be used to determine the approximate age of a healing hymenal laceration.

The configuration of a hymenal laceration included cleft-like patterns, whereby the wound edges remained relatively close together; V-shaped lacerations that had a sharp or pointed base; U-shaped configurations whose base was narrow but rounded; concavities whose base was both broad and rounded; and lacerations with a broad base and a narrow rim. The “healed” laceration category incorporated findings that could no longer be classified as a laceration. This included the presence of new blood vessel formation (neovascularity) and scar tissue formation at the former location of a laceration.

**Interobserver Reliability**

Individually, we performed a blinded reexamination of a random sample of 10% \((n = 25)\) of the cases to assess and measure the reliability of the original agreement on the interpretation of a finding. \(\kappa\) statistics were used to determine this interobserver reliability. The \(\kappa\) scores ranged from 0.46 to 1.0 (moderate to excellent). On the basis of the interpretation of \(\kappa\) results by Landis and Koch\(^6\) as well as Fleiss,\(^17\) the authors concluded that the results from this study are sufficiently reliable.

**Statistical Analysis**

We entered all collected data into an Access database. We documented the photographic findings on a worksheet and systematically entered it into the created database. These data were then transferred directly into SPSS (SPSS, Chicago, IL) for analysis. Descriptive statistics were used to show the results of each of the 3 examination methods. Data were analyzed using \(t\) tests, \(\chi^2\), Yates continuity correction or Fisher’s exact tests, and Mann-Whitney \(U\) tests, when appropriate. Statistical significance was defined as \(P < .05\).

**RESULTS**

The patients consisted of 239 female children and adolescents from 4 months to 18 years of age. There were 113 (47%) prepubertal girls and 126 (53%) adolescent girls. Of the 113 prepubertal children, 50% were white, 16% were black, 26% were Hispanic, 4% were Asian, and 4% were of a mixed race. Of the 126 pubertal girls, 48% were white, 29% were black, 14% were Hispanic, 4% were Asian, and 6% were of a mixed race.

**Timing of Examinations and the Cause of Injuries**

The period between an injury and the initial examination ranged from 1 hour to 3 days. A total of 164 (69%) of the 239 patients were seen within 24 hours after their injury. A total of 208 (87%) were examined within 48 hours. The other 31 (13%) girls were first evaluated between 48 and 72 hours after their injury. The mean time between an injury and the first examination was 24 hours. The causes of the injuries as determined by the contributors of the 113 prepubertal girls included 21 (19%) accidental or noninflicted injuries, 73 (65%) injuries secondary to abuse, and 19 (17%) “unknown cause” injuries. All 126 pubertal adolescents were said to be victims of a sexual assault.

**Summary of the Findings**

Because of the nature of this study, the timing of both the initial and the follow-up examinations varied as a result of the circumstances of each case. During each follow-up examination, the number of days since the
injury and the status of each hymenal abrasion, contusion, or laceration were recorded. The 113 prepubertal girls had 201 hymenal abrasions and contusions. The soonest “reevaluation” of a girl occurred within 24 hours of her initial examination. The longest a prepubertal girl was followed was 2.5 years. The average follow-up period was 9.9 months. The longest a prepubertal girl was followed was 2.5 years. The average follow-up period was 9.9 months. The average follow-up period was 61 days. The 126 adolescents were found to have 230 hymenal abrasions and contusions. The period for a reevaluation after an assault ranged from 1 day to 3.7 months. The average follow-up period was 61 days.

Table 1 is a compilation of the period required for a hymenal abrasion or contusion to resolve. The healing process was recorded as follows:

- **“Last detected”** identifies the day in which a finding was last detected in patients with a particular injury.
- **“Earliest disappearance”** identifies the day in which a particular finding was no longer identified in any one patient.
- **“Gone”** identifies the first examination day in which a finding was no longer seen in any of the patients.
- **“Never seen”** represents a finding that was never seen in any of the patients during a follow-up examination. Unfortunately, in this case there was no way of knowing when such an injury had actually disappeared.

For example, of the 8 (7%) prepubertal girls with a hymenal abrasion (for which there were a total of 13 follow-up examinations), the only time this finding was seen on a follow-up examination in any of the patients was on the day following the initial evaluation (day 1) (see “Last Detected” column). No abrasions were detected in any of the other follow-up examinations, beginning with 2 girls who were reexamined on day 3 (see “Earliest Disappearance” column). Day 3 was also the day after which no other hymenal abrasions were detected (see “Gone” column).

Hematoma is used as an example of the term “never seen” in Table 1. Five prepubertal girls had what initially appeared to be a hymenal “hematoma.” The soonest any of these girls were reexamined was 2 days after their injury (day 2). At that time, as well as on all of the other follow-up examinations, the well-defined, localized collection of blood (hematoma) on their hymens had been replaced by diffuse submucosal hemorrhages. Therefore, a hematoma was “never seen” after the initial examination.

### Hymenal Abrasions

See the previous example for the hymenal abrasions that were detected in the prepubertal girls. Only 2 (1%) adolescents had hymenal abrasions (Table 1). Their first reevaluation occurred on day 4, and, in both cases, the abrasions had disappeared, leaving only a localized area of erythema.

### Hymenal Contusions

#### Blood Blisters

The thin vesicles of blood (blood blisters) on the surface of the hymen were associated with the more severely injured patients. Once formed, this small, blood-filled vesicle seemed to shrink in size before disappearing completely.

Only 1 prepubertal girl was discovered to have a hymenal blood blister (Table 1). Although the blood blister was present on the seventh day after the injury, it was gone by the 11th day of follow-up.

### Table 1: Healing of Hymenal Abrasions and Contusions: Prepubertal and Pubertal Girls

<table>
<thead>
<tr>
<th>Type of Injury</th>
<th>Group</th>
<th>Severity</th>
<th>Last Detected</th>
<th>Earliest Disappearance</th>
<th>Gone</th>
<th>Never Seen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abrasions</strong></td>
<td>Pre: 8 (7%), F/U: 13</td>
<td></td>
<td>1 d (1/1)</td>
<td>3 d (2/2)</td>
<td>3–22 d (10/10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub: 2 (1%), F/U: 4</td>
<td></td>
<td>4 d (2/2)</td>
<td>11 d (2/2)</td>
<td>11 d (2/2)</td>
<td></td>
</tr>
<tr>
<td><strong>Blood blister</strong></td>
<td>Pre: 1 (1%), F/U: 3</td>
<td></td>
<td>7 d (1/1)</td>
<td>30 d (1/1)</td>
<td>Unknown (1/1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub: 7 (5%), F/U: 11</td>
<td></td>
<td>34 d (1/5)</td>
<td>9 d (1/5)</td>
<td>Unknown (1/1)</td>
<td></td>
</tr>
<tr>
<td><strong>Hematoma</strong></td>
<td>Pre: 5 (4%), F/U: 5</td>
<td></td>
<td>1 d (2/2)</td>
<td>3 d (2/2)</td>
<td>3–20 d (13/13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub: 13 (10%), F/U: 19</td>
<td></td>
<td>1 d (1/1)</td>
<td>3 d (1/1)</td>
<td>3–10 d (9/9)</td>
<td></td>
</tr>
<tr>
<td><strong>Petechiae</strong></td>
<td>Pre: 69 (60%), F/U: 97</td>
<td></td>
<td>2 d (1/2)</td>
<td>1 d (1/2)</td>
<td>2–21 d (90/90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub: 65 (50%), F/U: 80</td>
<td></td>
<td>3 d (1/7)</td>
<td>2 d (1/1)</td>
<td>3–10 d (72/72)</td>
<td></td>
</tr>
<tr>
<td><strong>Submucosal hemorrhages</strong></td>
<td>Pre: 51 (45%), 118 submucosal hemorrhages</td>
<td>Mild: 8, F/U: 13</td>
<td>2 d (3/3)</td>
<td>3 d (1/1)</td>
<td>3–94 d (10/10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mod: 53, F/U: 70</td>
<td>8 d (1/34)</td>
<td>5 d (2/8)</td>
<td>10–304 d (28/28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mark: 57, F/U: 94</td>
<td>15 d (1/54)</td>
<td>5 d (3/16)</td>
<td>16–730 d (22/22)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pub: 67 (53%), 137 submucosal hemorrhages</td>
<td>Mild: 16, F/U: 22</td>
<td>7 d (1/8)</td>
<td>4 d (4/4)</td>
<td>8–29 d (10/10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mod: 71, F/U: 80</td>
<td>12 d (2/48)</td>
<td>3 d (1/5)</td>
<td>14–36 d (37/37)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mark: 50, F/U: 66</td>
<td>11 d (1/23)</td>
<td>3 d (1/5)</td>
<td>12–97 d (38/38)</td>
<td></td>
</tr>
</tbody>
</table>

Values in parentheses represent the number of patients with a particular finding on any given day. Pre indicates prepubertal girls (nonestrogenized hymen); Pub, pubertal adolescents (estrogenized hymen); F/U, follow-up examinations; Mod, moderated; Mark, marked.

- a The day in which a finding was last seen during follow-up in all patients with this finding.
- b The first day in which a finding was no longer identified in any one patient.
- c The first follow-up day in which a finding was no longer seen in any of the patients with this finding.
- d A finding is never seen in any follow-up of all of the patients with this finding.
the actual time of resolution is unknown because she did not return for any additional reevaluations.

Seven (5%) adolescents had blood blisters on their hymenal membranes (Table 1). These lesions were detected for the first time in 5 patients during the second and third postinjury weeks. One adolescent still had a blood blister on day 34. There were no additional examinations of this group of adolescents.

**Erythema**

The redness of the tissues that is created by capillary congestion (erythema) constitutes a nonspecific finding. Therefore, erythema is not included as a variable in “Results” because of its uncertain clinical significance.

**Hematomas**

What initially seemed to be a well-defined, localized collection of blood (hematoma) dramatically changed during a relatively short period as the blood disseminated into the surrounding tissues (Table 1). At that point, they were considered to be submucosal hemorrhages.

**Petechiae**

Sixty-nine (60%) of the 113 prepubertal girls had a pinpoint, nonraised, perfectly round, purplish red spot (petechia) on their hymenal membranes at the time of their initial examination (Table 1). No petechiae were detected beyond 48 hours in any of the prepubertal girls. Sixty-five (50%) of the 126 adolescents had petechiae on their hymens at the time of their initial evaluation (Table 1). No petechiae were identified in any of these pubertal girls after 72 hours.

**Submucosal Hemorrhages**

Submucosal hemorrhages were discovered in 51 (45%) of the 113 prepubertal girls and in 67 (53%) of the 126 pubertal adolescents. Evidence of this bleeding into the areolar tissue beneath the mucosal membrane was found primarily in the posterior quadrants of the hymen in both age groups. The depth of discoloration of a submucosal hemorrhage and its relative size in relationship to the surrounding tissue was used in classifying them as mild, moderate, or marked. Each lesion was individually tracked, and the disappearance day was recorded. The more severe hemorrhages gradually evolved into either a moderate or mild form before completely disappearing (Table 1).

**Hymenal Lacerations**

The 40 hymenal lacerations that were observed in the 113 prepubertal girls were reevaluated a total of 60 times. The 80 hymenal lacerations that were identified in the 126 pubertal adolescents were reexamined a total of 93 times. The locations of these lacerations were recorded in relationship to the face of a clock as though the patient were in a supine position. As the hymenal lacerations healed, several changes took place. These included variations in both the depth and the configuration of the laceration.

The location of the hymenal lacerations varied somewhat by age (Table 2). Both groups of patients had significantly more \( P < .01 \) lacerations on the posterior half of their hymenal rim than on the anterior portion of this membrane. Of the posterior rim lacerations, 75% of the prepubertal girls’ lacerations were in or close to the midline, whereas only 29% of the adolescents’ lacerations were found at this same area \( P < .001 \). Conversely, the older patients had a greater percentage of lacerations along the lateral hymenal rim at the 3 o’clock and 9 o’clock locations \( P < .05 \) than the younger girls.

**Depth of the Healing Hymenal Lacerations**

**Prepubertal Girls**

The depth of the hymenal lacerations in the prepubertal girls ranged from superficial tears to transactions that extended into the fossa navicularis and beyond.

<table>
<thead>
<tr>
<th>Location</th>
<th>Prepubertal</th>
<th>Pubertal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior</td>
<td>2 5</td>
<td>12 15</td>
</tr>
<tr>
<td>Lateral</td>
<td>3 8</td>
<td>18 23</td>
</tr>
<tr>
<td>Posterior</td>
<td>35 88</td>
<td>50 61</td>
</tr>
<tr>
<td>Total</td>
<td>40 100</td>
<td>80 100</td>
</tr>
</tbody>
</table>

* Significant at \( P < .01 \).

**Table 2** Location of Hymenal Lacerations: Comparison of Prepubertal and Pubertal Girls

<table>
<thead>
<tr>
<th>Classification</th>
<th>Contusion, ( n (%) )</th>
<th>Superficial, ( n (%) )</th>
<th>Intermediate, ( n (%) )</th>
<th>Deep, ( n (%) )</th>
<th>Transection, ( n (%) )</th>
<th>Transection With an Extension, ( n (%) )</th>
<th>Undetected or Healed, ( n (%) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2: Superficial</td>
<td>1 (20)</td>
<td>1 (20)</td>
<td>1 (20)</td>
<td>2 (40)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3: Intermediate</td>
<td>2 (33)</td>
<td>1 (17)</td>
<td>1 (17)</td>
<td>2 (33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4: Deep</td>
<td>2 (10)</td>
<td>3 (14)</td>
<td>13 (61)</td>
<td>2 (10)</td>
<td>1 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5: Transection</td>
<td>1 (14)</td>
<td>2 (29)</td>
<td>3 (43)</td>
<td></td>
<td>1 (14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6: Transection With an extension</td>
<td>7 (23)</td>
<td>14 (45)</td>
<td>9 (29)</td>
<td></td>
<td>1 (3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F/U indicates number of acute lacerations per number of follow-up examinations.
(Table 3). As they healed, the apparent depth of a number of the lacerations changed. Whereas most of the lacerations took on a more superficial appearance, others became deeper in appearance as the swelling of the surrounding tissues receded. Sixty-eight percent of transections with an extension evolved into either transections without an extension or into deep-appearing lacerations as they healed. The reverse of this trend was also observed. Fifteen percent of deep lacerations turned out to be transections in which all evidence of a hymenal rim at the base of the laceration disappeared as the swelling of the tissues subsided.

The case of an 8-month-old is an example of the healing process that was observed in several of the prepubertal girls. Her hymenal (Fig 1, arrow), fossa navicularis, and posterior fourchette lacerations, as seen on the third day after the assault, dramatically changed in appearance during the subsequent month. By the 28th day, the posterior rim of her hymen appears smooth and relatively wide (Fig 2). Her fossa navicularis and posterior fourchette injuries have disappeared, leaving only a collection of small blood vessels (neovascularity) at the site of her injuries.

**Pubertal Adolescents**

The recorded depth of the hymenal lacerations also changed in many of the older adolescents as their wounds healed (Table 4). As the swelling of the tissues subsided and the submucosal hemorrhages disappeared, a number of the wounds took on a shallower appearance. In contrast, this same phenomenon also exposed deeper injuries. Lacerations in 2 adolescents were originally classified as deep. As the swelling of their tissues subsided, the lacerations had to be reclassified as transections. During a follow-up examination, evidence of a laceration was no longer detectable in 3 adolescents because of the redundancy of their hymenal tissues.

The case of a 14-year-old adolescent exemplifies how the presence of edema and submucosal hemorrhage can effect the findings (Figs 3 and 4). This girl was sexually assaulted 12 hours before her initial examination. By the fourth day after the assault, the swelling had receded and most of the submucosal hemorrhage had disap-
peared (Figs 5 and 6). In the prone, knee-to-chest position (Fig 6), the multiple lacerations of her hymen became apparent and her hymenal orifice took on a “starburst” appearance. On the 16th day after injury (Fig 7), her evaluation by the supine, labial traction technique revealed only a single “cleft” at the 8 o’clock position. As she was repositioned into the prone, knee-to-chest position, evidence of her multiple lacerations once again became apparent (Fig 8).

### Changes in Configuration of a Healing Hymenal Laceration

**Prepubertal Girls**

The most common configuration of the acute hymenal lacerations in the prepubertal age group was a V-shaped laceration (39%). Of the 40 acute lacerations that were identified in the prepubertal girls, 21% were cleft-like. On healing, the proportion of clefts decreased to 10%. The percentage of V-shaped lacerations dropped from 39% to 18%. The smoother, U-shaped configurations increased from 21% to 46% as they healed. The

<table>
<thead>
<tr>
<th>Classification</th>
<th>Superficial, n (%)</th>
<th>Intermediate, n (%)</th>
<th>Deep, n (%)</th>
<th>Transection, n (%)</th>
<th>Transection With an Extension, n (%)</th>
<th>Undetected or Healed, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2: Superficial (13/14 F/U)</td>
<td>8 (57)</td>
<td>3 (21)</td>
<td>1 (7)</td>
<td></td>
<td></td>
<td>2 (15)</td>
</tr>
<tr>
<td>3: Intermediate (9/12 F/U)</td>
<td>3 (25)</td>
<td>9 (75)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4: Deep (22/24 F/U)</td>
<td>5 (21)</td>
<td>15 (63)</td>
<td>2 (8)</td>
<td></td>
<td></td>
<td>1 (4)</td>
</tr>
<tr>
<td>5: Transection (23/26 F/U)</td>
<td>2 (8)</td>
<td>3 (12)</td>
<td>12 (46)</td>
<td>7 (27)</td>
<td></td>
<td>2 (8)</td>
</tr>
<tr>
<td>6: Transection with an extension (13/17 F/U)</td>
<td>2 (12)</td>
<td>3 (18)</td>
<td>10 (59)</td>
<td>2 (12)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
broad-based, U-shaped lacerations increased slightly from 18% to 23%. On healing, 3% of the prepubertal girls had perfectly smooth, wide hymenal rims that revealed no evidence of the previous trauma. The case of the 9-year-old whose father confessed to having sexually assaulted his daughter 3 days before (Fig 9) is an example of a V-shaped laceration’s smoothing off and taking on a smoother, “key-hole” type of appearance (Fig 10) by the 19th day after assault.

**Pubertal Adolescents**

The most common configuration of the acute hymenal laceration in the adolescent was also V-shaped (48%). Of the 80 acute lacerations that were identified in the adolescent girls, 28% were cleft-like. The percentage of these cleft-like lacerations decreased to 22% as they healed. The percentage of V-shaped lacerations decreased from 48% to 24% over time. The smoother, narrow, U-shaped configurations increased from 24% to 34%, whereas the broad-based, U-shaped–appearing lacerations increased significantly from 1% to 20% ($P < .05$). Three percent of the hymenal lacerations could no longer be identified during a follow-up examination.

**Outcome of the Hymenal Lacerations**

The healing process of a hymenal injury varied with the extent of the injury. Evidence of a “recent” injury faded rapidly. This included the disappearance of edema, petechiae, submucosal hemorrhages, and fresh-cut surfaces. Beyond this initial period, the lacerations continued to undergo changes in both depth and configuration for up to 3 and 4 weeks. The sites of the healed lacerations varied in smoothness, continuity, and width (Tables 5 and 6). No scar tissue was identified on the hymen of any of the girls.

**Prepubertal Girls**

No evidence of an acute hymenal laceration was present by the end of the first week in 2 of the 4 prepubertal girls who were examined during that period. By 10 days, 9 (82%) of 11 patients no longer had signs of an acute injury. Except for the 1 girl with the blood blister, all evidence of an acute hymenal injury had
disappeared by the 16th (22 of 22) day after injury. No change in either the depth or the configuration of a healing laceration was observed beyond 30 days.

The appearance of the healed lacerations on the hymenal rims of 2 prepubertal girls remained constant during the 2 and 3 years they were followed. Evidence of a laceration in 2 other prepubertal girls disappeared into the folds of their estrogenized hymens when they entered into puberty.

The final outcome of the prepubertal girl’s hymenal laceration was dictated by the extent of the injury (Table 5). When the results of the superficial, intermediate, and deep lacerations were combined, 75% (15 of 20) of the prepubertal girls had smooth hymenal rims with no disruption in contour (“continuous”). Of those who had sustained a transection or a transection with an extension, 17% (3 of 18) had a smooth rim, whereas 22% (4 of 18) had a continuous-appearing hymenal membrane on healing. The hymenal rim widths measured <1 mm in 28% (5 of 18) of the girls who had sustained either a transection or a transection with an extension. Hymenal rim width measurements were not obtained in the girls with the less severe injuries.

**Pubertal Adolescents**

Evidence of a recent injury disappeared in the adolescents at approximately the same rate as their prepubertal counterparts. In the first 7 days, 5 (56%) of 9 adolescents no longer had signs of edema, erythema, submucosal hemorrhage, or fresh-cut surfaces. At 10 days, 9 of 9 still had evidence of an acute injury. By 2 weeks, 90% (19 of 21) of the lacerations appeared healed. By 3 weeks, only those with blood blisters still had signs of a recent injury. None of the adolescents were followed for >90 days.

Similar to the findings in the prepubertal girls, the final outcome of an adolescent’s hymenal laceration was determined by the extent of the injury (Table 6). When the results of superficial, intermediate, and deep lacerations were combined, 59% (24 of 41) of the hymenal rims had a normal, scalloped appearance and 88% (30 of 34) revealed no disruption in continuity. Thirty-eight percent of the hymenal rims of pubertal adolescents who
sustained either a transection or a transection with an extension had a narrow but normal, scalloped appearance. Fifty-nine percent (17 of 29) revealed no disruption of the contour of their hymenal membrane. Eighty-seven percent (21 of 24) did have rims that measured <1 mm in width. No hymenal scars were identified in any of the prepubertal or adolescent girls.

DISCUSSION
Many of the findings in this study are similar to those in other reports on hymenal injuries. The majority of the hymenal lacerations in this study did smooth off over time, as previously reported by McCann et al. The variety in the configuration of the healed lacerations described by Kerns et al is similar to those observed in this report. Although the terminology differed, the outcome of the 37 girls with hymenal injuries in the report by Heppenstall-Heger et al seems to be similar to that found in this study. One difference in the results of their study with ours is the location of the hymenal lacerations. Heppenstall-Heger et al reported, “All tears occurred in the posterior 180 degrees, between 4 and 8 o’clock, except for 1 accidental avulsion injury.” We identified hymenal lacerations at all locations on the hymenal rim in both the prepubertal and the adolescent girls (Table 2). One possible explanation is the type of examination methods used. Several of the anterior and lateral lacerations that we identified were detected only during the prone, knee-to-chest position. Although Heppenstall-Heger et al reported that they used this method, it was used only “when abnormalities were noted.”

The observation by Berenson et al of a hymenal rim width of <1 mm in their 3- to 8-year-old girls who had “a history of penetration” was consistent with the outcome of some but not all of the girls in this study. In our report, 13 (72%) of 18 of the prepubertal girls who had sustained a laceration that either transected the hymen or extended through the hymenal attachment and into the surrounding tissues still had a hymenal rim width of >1 mm on healing (Table 5). This phenomenon in the prepubertal girl was attributed to the development of a very thin, delicate membrane that appeared at the base.
of the laceration as the healing took place. In the adolescents, the width of the healed hymenal rim was >1 mm in 13% after the healing of these 2 deeper types of lacerations (Table 6).

The commentary by Adams15 on the likelihood of a finding being normal if there is a “continuity of the hymenal rim” seems reasonable. Unfortunately, her comment did not take into account the remarkable healing process of the injured hymenal membrane. In our study, the majority of the prepubertal girls still had a smooth edge and continuity of the hymenal rim after the healing of these 2 deeper types of lacerations (Table 5).

The second marker, at the other end of the spectrum, was the presence of a blood blister (Fig 8). These small, blood-filled vesicular lesions, which frequently appeared for the first time during a follow-up examination, indicated that an injury had occurred sometime in the past month. This marker was particularly helpful in the adolescent cases when all other signs of an acute injury had disappeared.

Completion of the healing process was defined by the disappearance of the signs of an acute injury and the cessation of changes in the depth and the configuration of a laceration. Whereas most signs of an acute injury were gone within 7 to 10 days, the changes in the depth and the configuration of a laceration continued for up to 3 weeks in the prepubertal girl and 4 weeks in the adolescent girl.

In this study, the findings that were created by a hymenal laceration in 2 prepubertal girls remained unchanged until they reached puberty. The findings in 2 other girls disappeared into the folds of their estrogenized hymens as they entered into puberty. None of the

### TABLE 5 Healed Hymenal Rim Findings by Depth of Laceration: Prepubertal Girls

<table>
<thead>
<tr>
<th>Depth</th>
<th>Yes, n (%)</th>
<th>No, n (%)</th>
<th>UTD, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superficial (n = 4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smooth appearance</td>
<td>2 (50)</td>
<td>1 (25)</td>
<td>1 (25)</td>
</tr>
<tr>
<td>Continuous</td>
<td>2 (50)</td>
<td>1 (25)</td>
<td>1 (25)</td>
</tr>
<tr>
<td>Delicate</td>
<td>1 (25)</td>
<td>2 (50)</td>
<td>1 (25)</td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>1 (25)</td>
<td>3 (75)</td>
<td></td>
</tr>
<tr>
<td>Intermediate (n = 4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smooth appearance</td>
<td>4 (100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>4 (100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delicate</td>
<td>2 (50)</td>
<td>2 (50)</td>
<td></td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>1 (25)</td>
<td>3 (75)</td>
<td></td>
</tr>
<tr>
<td>Deep (n = 11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smooth appearance</td>
<td>9 (82)</td>
<td>2 (18)</td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>9 (82)</td>
<td>2 (18)</td>
<td></td>
</tr>
<tr>
<td>Delicate</td>
<td>4 (36)</td>
<td>7 (64)</td>
<td></td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>1 (9)</td>
<td>10 (91)</td>
<td></td>
</tr>
<tr>
<td>Transection (n = 4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smooth appearance</td>
<td>2 (50)</td>
<td>1 (25)</td>
<td>1 (25)</td>
</tr>
<tr>
<td>Continuous</td>
<td>2 (50)</td>
<td>1 (25)</td>
<td>1 (25)</td>
</tr>
<tr>
<td>Delicate</td>
<td>1 (25)</td>
<td>2 (50)</td>
<td>1 (25)</td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>1 (25)</td>
<td>2 (50)</td>
<td></td>
</tr>
<tr>
<td>Transection/extension (n = 16)</td>
<td>1 (6)</td>
<td>14 (88)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Smooth appearance</td>
<td>2 (13)</td>
<td>13 (81)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Continuous</td>
<td>2 (13)</td>
<td>13 (81)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Delicate</td>
<td>1 (6)</td>
<td>15 (94)</td>
<td></td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>4 (25)</td>
<td>12 (75)</td>
<td></td>
</tr>
</tbody>
</table>

UTD indicates unable to determine for a variety of reasons.

### TABLE 6 Healed Hymenal Rim Findings by Depth of Laceration: Pubertal Girls

<table>
<thead>
<tr>
<th>Depth</th>
<th>Yes, n (%)</th>
<th>No, n (%)</th>
<th>UTD, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superficial (n = 13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal scalloped appearance</td>
<td>8 (62)</td>
<td>4 (31)</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Continuous</td>
<td>11 (50)</td>
<td>1 (8)</td>
<td>1 (8)</td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>3 (23)</td>
<td>6 (46)</td>
<td>4 (31)</td>
</tr>
<tr>
<td>Intermediate (n = 9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal scalloped appearance</td>
<td>6 (67)</td>
<td>2 (22)</td>
<td>1 (11)</td>
</tr>
<tr>
<td>Continuous</td>
<td>7 (78)</td>
<td>1 (11)</td>
<td>1 (50)</td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>3 (33)</td>
<td>3 (33)</td>
<td>3 (33)</td>
</tr>
<tr>
<td>Deep (n = 20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal scalloped appearance</td>
<td>10 (46)</td>
<td>11 (50)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Continuous</td>
<td>12 (60)</td>
<td>2 (9)</td>
<td>5 (23)</td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>10 (46)</td>
<td>6 (27)</td>
<td>6 (27)</td>
</tr>
<tr>
<td>Transection (n = 23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal scalloped appearance</td>
<td>8 (35)</td>
<td>12 (52)</td>
<td>3 (13)</td>
</tr>
<tr>
<td>Continuous</td>
<td>11 (48)</td>
<td>8 (35)</td>
<td>4 (17)</td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>14 (61)</td>
<td>1 (4)</td>
<td>8 (45)</td>
</tr>
<tr>
<td>Transection/extension (n = 13)</td>
<td>4 (31)</td>
<td>8 (62)</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Normal scalloped appearance</td>
<td>6 (46)</td>
<td>5 (39)</td>
<td>2 (15)</td>
</tr>
<tr>
<td>Continuous</td>
<td>6 (46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 mm in width</td>
<td>7 (54)</td>
<td>2 (15)</td>
<td>4 (31)</td>
</tr>
</tbody>
</table>

UTD indicates unable to determine for a variety of reasons.
adolescents was followed long enough to determine how their findings might have changed over time. The data from this study did not reveal any difference in the healing process between the prepubertal and pubertal girls. The nonhymenal genital injuries data on the healing process and their outcome in this population of girls will be presented in a companion report.

CONCLUSIONS
The hymenal injuries in these prepubertal and adolescent girls all healed rapidly and frequently left little or no evidence of the previous trauma. The rapid resolution of the petechiae along with the persistence of blood blisters did provide markers for approximating the age of an injury. The multiple locations of the lacerations on the hymenal rim are a reminder of the importance of the multimethod approach during these examinations. Although the outcome and the final appearance of a hymenal laceration depended on its severity, the smoothness and the persistent continuity of a hymenal rim after all but the most severe lacerations should prove to be reassuring to the victim and her family. No scar tissue was identified on the hymen in any of the patients. These findings reaffirm the remarkably complex healing process that occurs after a hymenal injury. These data heighten the examiner’s need to exercise caution before calling a finding “normal, without evidence of a previous injury.”

APPENDIX: GLOSSARY OF TERMS

1. Cleft: An angular defect on the edge of the hymen whose edges are closely approximated. The defect may extend to the muscular attachment of the hymen.
2. Concavity: A curved or hollowed U-shaped depression of the edge of the hymenal membrane.
3. Notch: A V-shaped indentation or defect on the edge of the hymenal membrane that may extend to the muscular attachment of the hymen.
4. Hymenal tear/laceration: A defect (injury) in the hymenal membrane caused by a blunt object that has ripped or pulled apart (rendering) the tissue.
5. Superficial partial tear of the hymenal membrane: A laceration or tear of the hymenal membrane that extends less than halfway through the width of the membrane.
6. Intermediate partial tear of the hymenal membrane: A laceration or tear of the hymenal membrane that extends halfway through the width of the membrane.
7. Deep partial tear of the hymenal membrane: A laceration or tear of the hymenal membrane that extends more than halfway through the width of the membrane.
8. Complete tear or transection of the hymen: A laceration or tear of the hymenal membrane that extends through the entire width of the membrane to its attachment.
9. Transection of the hymen with an extension: A laceration or tear of the hymenal membrane that extends through the attachment and into the surrounding tissues.
10. Laceration: A defect of the tissues caused by a ripping or pulling apart (rendering). The wound may contain bridging structures.
11. Incision: A wound created by a sharp instrument whose edges are well defined. The wound contains no bridging structures.

ACKNOWLEDGMENTS
This study was supported in part from the Lucile Packard Foundation; the Kenneth Jonsson Family Foundation; and the Children’s Miracle Network through the Department of Pediatrics, University of California, Davis.

We thank the following individuals for their participation in this multicenter, collaborative research project:

David Kerns, MD, Valley Medical Center, San Jose, California; Marilyn Kaufhold, MD, San Diego Children’s Hospital; Mary Ritter, CHA, Valley Medical Center, San Jose, California; Joan Voris, MD, University Hospital and Medical Center, Fresno, California; Terrence Donald, MD, Women’s and Children’s Hospital, North Adelaide, Australia; Martin Finkel, DO, University of Medicine and Dentistry of New Jersey, Stratford, New Jersey; Lori Frasier, MD, University Physicians-Green Meadows, Columbia, Missouri; Wendy Gladstone, MD, Exeter Pediatric Associates, Exeter, New Hampshire; Penny Grant, Child Abuse Network, Tulsa, Oklahoma; Nancy Harper, MD, Naval Medical Center Portsmouth, Virginia; Roberta Hibbard, MD, Indiana University School of Medicine, Indianapolis, Indiana; Ralph Hicks, MD, Wright State University, Dayton, Ohio; Margie Hogan, MD, Hennepin County Medical Center, Minneapolis, Minnesota; Michael Jordan, MD, Medical Office Building, Newark, New York; Michael Knappman, PA-C, Redwood Children’s Center, Santa Rosa, California; Kathe Kraley, RN, MSN, CPNP, Children’s Mercy Hospital, Kansas City, Missouri; Susan Murawski, MS, ARNP, CPNP, Southwest Florida Children’s Fund, Fort Myers, Florida; Vincent Palusci, MD, Michigan State University College of Human Medicine, Grand Rapids, Michigan; Andrew Sirotnak, MD, Children’s Hospital, Denver, Colorado; Naomi Sugar, MD, HCSATS, Seattle, Washington; Andi Taroli, MD, Children’s Advocacy Center, Scranton, Pennsylvania; Danny Waldrop, MD, Geisinger Medical
Center, Danville, Pennsylvania; J.M. Whitworth, MD, Florida Children’s Fund, Jacksonville, Florida.

Special thanks go to Ana Ross, PA-C, for assistance in reviewing the medical charts of the UCDMC patients; to Stephen Boos, MD, Lieutenant Colonel, United States Air Force Medical Corps, for review of the manuscript; and to Lieschen Trelford for assistance with the myriad of tasks connected to the preparation of this manuscript.

REFERENCES

15. Adams J. Normal studies are essential for objective medical evaluation of children who may have been sexually abused. Acta Paediatr. 2003;92:1378–1380
The Effect of Breastfeeding on Cardiorespiratory Risk Factors in Adult Life

Alicja R. Rudnicka, PhD, Christopher G. Owen, PhD, David P. Strachan, MD

Division of Community Health Sciences, St George’s, University of London, London, United Kingdom

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

ABSTRACT

OBJECTIVE. Nutrition in the first weeks of life may program disease risk in adulthood. We examined the influence of initial infant feeding on cardiorespiratory risk factors in adulthood.

PATIENTS AND METHODS. A total of 9377 persons born during 1 week in 1958 in England, Scotland, and Wales were followed-up periodically from birth into adulthood. Infant feeding was recorded from a parental questionnaire at 7 years old as never breastfed, breastfed partially or wholly for <1 month, or breastfed for >1 month. Height; waist circumference; hip circumference; waist/hip ratio; body mass index; blood pressure; forced expiratory volume; total, high-density, and low-density lipoprotein cholesterol; triglycerides; hemoglobin A1c; fibrinogen; fibrin D-dimer; C-reactive protein; von Willebrand factor; and tissue plasminogen activator antigen were measured at 44 to 45 years of age.

RESULTS. Breastfeeding for >1 month was associated with reduced waist circumference, waist/hip ratio, von Willebrand factor, and lower odds of obesity compared with formula feeding after adjustment for birth weight, prepregnancy maternal weight, maternal smoking during pregnancy, socioeconomic position in childhood and adulthood, region of birth, gender, and current smoking status. Infant feeding status was not associated with other cardiorespiratory risk factors after adjustment, except for lower fibrinogen and C-reactive protein levels in women.

CONCLUSIONS. The inverse associations of breastfeeding for >1 month with measures of central obesity and inflammatory markers in the current study are small and of little public health importance. Although there was no substantial long-term protective effect of breastfeeding for >1 month on other cardiorespiratory risk factors in adult life, further studies with contemporaneous data on exclusive breastfeeding are needed to confirm these findings.
CARDIORESPIRATORY DISORDERS are a major cause of death in later life in both the developed and developing world. It has been suggested that nutrition in the first weeks of life may program disease risk in adulthood. Any such effect is of considerable public health interest, because early infant feeding patterns are potentially modifiable. Breast milk has long been encouraged as the food of choice in infancy, based on numerous short- and long-term health benefits. Breastfeeding has been associated with an improved cardiovascular risk profile and fewer cardiovascular outcomes in later life (including ischemic heart disease and type 2 diabetes) compared with formula feeding. However, there is insufficient evidence to conclude whether breastfeeding is protective against cardiovascular mortality. It has also been suggested that initial breastfeeding may result in taller height and reduce the prevalence of obesity. However, the results of individual studies for this latter association have differed, showing either protective or null effects. Evidence that breastfeeding influences lung function and respiratory outcomes in later life has been limited and inconclusive.

Recent systematic reviews have highlighted that publication bias (where smaller studies showing beneficial effects of breastfeeding on health outcome are preferentially published) and lack of adjustments for potentially important confounders are key concerns in establishing whether apparent associations between breastfeeding and health outcomes are causally related. It is also difficult to examine the long-term consistency of these associations, because few studies have contemporary measurement of infant feeding with follow-up in adult life. This study aims to examine the association between the mode of infant feeding and a range of established cardiorespiratory risk factors in adulthood in a large prospective cohort study, where data on infant feeding were ascertained in childhood and information on a large number of potentially important confounders were measured throughout the life course.

METHODS

Study Design

Persons born in England, Scotland, and Wales during 1 week in March 1958 were recruited for a Perinatal Mortality Survey and have been followed up periodically from birth into adulthood (the British 1958 birth cohort). Participants were invited to a clinical examination by a trained research nurse visiting their home during 2002 to 2004, when cohort members were aged 44 to 45 years. From a target sample of 12 069 persons, 9377 (78%) were visited in their home by a team of 122 specially trained nurses from the National Centre for Social Research, who conduct the annual health surveys of England and Scotland.

Infant feeding was recalled by parents (usually the mother) when the child was 7 years of age into 1 of 4 categories: breastfed wholly or partially for >1 month, <1 month, not at all, and unknown. The following were identified as potentially important confounders: region of birth, prepregnancy maternal body size, maternal smoking during pregnancy, birth weight, and socioeconomic position in childhood and adulthood. Region of birth, prepregnancy maternal body size, maternal smoking during pregnancy, and birth weight were obtained from the 1958 Perinatal Mortality Survey. Socioeconomic position in childhood was based on father’s occupation in 1958 or at 7 years of age if data were unavailable at birth; adult social class was based on the participant’s current or most recent occupation at 42 years (or 33 years, if missing). Current smoking status (at 42 years) was defined as never smoker, ex-smoker, or current smoker. Social position and smoking status in adulthood are considered as potential later-life confounders, because they are important predictors of cardiorespiratory risk and are related to infant feeding status in this cohort.

Measurements and Assays in Adulthood

Standing height (measured using a Leicester portable stadiometer [Child Growth Organization, London, United Kingdom]), waist, and hip circumference were measured to the nearest millimeter. Weight was measured to the nearest 0.1 kg in light clothing with shoes removed. BMI was determined as kilograms per meter squared. Blood pressure and pulse rate were measured 3 times (mean used for analyses) in the seated position after a period of 5 minutes of rest using the Omron 705CP automated sphygmomanometer (Omron, Tokyo, Japan), with a large cuff for subjects with a mid-upper arm circumference of ≥32 cm. Spirometry was performed in the standing position without nose clips, using the Vitalograph Micro spirometer (Vitalograph Ltd, Buckingham, England). Forced expiratory volume in 1 second (FEV1) was standardized for current height.

Nonfasting venous blood samples were drawn into Sarstedt (Leicester, United Kingdom) polypropylene tubes containing citrate anticoagulant (for measurement of glycated hemoglobin [HbA1c], fibrinogen, tissue plasminogen activator antigen [t-PA], and von Willebrand factor [vWF] antigen) and no anticoagulant (for measurement of total cholesterol and high-density lipoprotein [HDL] cholesterol and triglycerides). Total and HDL cholesterol were measured by autoanalyzer, and low-density lipoprotein (LDL) cholesterol levels were derived by algorithm. Fibrinogen was determined by the Clauss method (MDA 180 coagulometer; Biomerieux, Basingstoke, United Kingdom). Fibrin D-dimer was measured by ELISA assay (Hyphen, Paris, France). C-reactive protein (CRP) was measured by high-sensitivity nephelometric analysis of latex particles coated with CRP monoclonal antibodies (BN ProSpec protein ana-
lyzer; Dade Behring, Marburg, Germany). T-PA and vWF were measured by ELISA assay using a double antibody sandwich (from Biopool, Umeå, Sweden, and DAKO, Copenhagen, Denmark, respectively). Written informed consent was obtained from all of the study participants. Ethical approval for the medical examination at 44 to 45 years was obtained from the South East Multicenter Research and Ethnics Committee (ref: 01/1/44).

Statistical Analysis
Waist circumference, hip circumference, waist/hip ratio, HDL cholesterol, triglycerides, HbA1c, fibrinogen, D-dimer, CRP, t-PA, and vWF were log transformed to normalize their distributions. The association between each risk factor and infant feeding status was investigated using a test for trend across all of the infant feeding groups ignoring the group with the unknown infant feeding status. Linear regression was used to compare those breastfed for <1 month and >1 month with those not breastfed at all with adjustment for region of birth, prepregnancy maternal body size, maternal smoking during pregnancy, birth weight, socioeconomic position in childhood and adulthood, smoking in adulthood, and gender. Those with missing data on confounders were included as a separate category in the analyses. Based on previous findings of gender differences in the association between infant feeding and cardiovascular risk,23,24 regression models were extended to test for an interaction between infant feeding status and gender. Parallel analyses were undertaken using logistic regression to examine the relation between the odds of obesity (individuals with BMI $\geq$ 30 kg/m$^2$) and duration of breastfeeding. Differences between the infant feeding groups for the outcomes that were log transformed are reported as percentage differences rather than absolute differences on the log scale. All of the regression analyses included additional adjustment for factors contributing to measurement error of the cardiorespiratory risk factors, such as nurse and instrument effects and laboratory batch; these nuisance factors are specified for each outcome in the results. All of the analyses were performed using Stata 9.2 (Stata Corp, College Station, TX).

RESULTS
Of the 9377 subjects who participated in the clinical examination at 44 to 45 years, infant feeding status was available for 8172 (87%); the remainder are shown in Table 1 as “unknown” infant feeding status and are excluded from all of the regression analyses. Of those with data on infant feeding, 46% ($n = 3772$) were breastfed for >1 month, 24% ($n = 1960$) were breastfed for <1 month, and 30% ($n = 2440$) were not breastfed at all. Before adjustment, increasing the duration of breastfeeding showed a strong positive linear association with adult height and HDL cholesterol and an inverse trend with BMI, obesity, waist circumference, waist/hip ratio, fibrinogen, D-dimer, t-PA, and vWF. Dur-

### TABLE 1

Comparison of Cardiorespiratory Risk Factors in Adults Who Were Breastfed for Different Durations

<table>
<thead>
<tr>
<th>Cardiorespiratory Risk Factors in Adulthood</th>
<th>Not Knowna</th>
<th>Not Breastfed</th>
<th>Breastfed &lt;1 mo</th>
<th>Breastfed &gt;1 mo</th>
<th>$P$ for Trendb</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Arithmetic mean (SD) by infant feeding status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height, cm</td>
<td>169.0 (9.1)</td>
<td>168.9 (9.3)</td>
<td>168.9 (9.4)</td>
<td>169.7 (9.3)</td>
<td>0.2</td>
</tr>
<tr>
<td>BMI, kg/m$^2$</td>
<td>27.4 (5.1)</td>
<td>27.8 (5.2)</td>
<td>27.5 (5.0)</td>
<td>27.2 (5.0)</td>
<td>0.0005</td>
</tr>
<tr>
<td>Systolic BP, mm Hg</td>
<td>126.2 (16.5)</td>
<td>127.2 (16.4)</td>
<td>125.8 (15.8)</td>
<td>126.7 (16.9)</td>
<td>0.61</td>
</tr>
<tr>
<td>Diastolic BP, mm Hg</td>
<td>78.9 (10.8)</td>
<td>79.1 (10.6)</td>
<td>78.4 (10.4)</td>
<td>78.8 (11.1)</td>
<td>0.61</td>
</tr>
<tr>
<td>Total cholesterol, mmol/L</td>
<td>3.38 (1.05)</td>
<td>3.44 (1.09)</td>
<td>3.44 (1.11)</td>
<td>3.41 (1.09)</td>
<td>0.0001</td>
</tr>
<tr>
<td>LDL cholesterol, mmol/L</td>
<td>3.38 (0.92)</td>
<td>3.44 (0.90)</td>
<td>3.44 (0.92)</td>
<td>3.41 (0.90)</td>
<td>0.11</td>
</tr>
<tr>
<td>Height adjusted FEV$_1$, L</td>
<td>3.25 (0.52)</td>
<td>3.28 (0.52)</td>
<td>3.29 (0.51)</td>
<td>3.30 (0.52)</td>
<td>0.09</td>
</tr>
<tr>
<td><strong>Geometric mean (GSD) by infant feeding statusc</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waist circumference, cm</td>
<td>91.0 (1.2)</td>
<td>91.6 (1.2)</td>
<td>90.9 (1.2)</td>
<td>90.2 (1.2)</td>
<td>0.0003</td>
</tr>
<tr>
<td>Hip circumference, cm</td>
<td>105.0 (1.1)</td>
<td>105.4 (1.1)</td>
<td>105.0 (1.1)</td>
<td>104.8 (1.1)</td>
<td>0.06</td>
</tr>
<tr>
<td>Waist/hip ratio</td>
<td>0.87 (1.11)</td>
<td>0.87 (1.10)</td>
<td>0.87 (1.10)</td>
<td>0.86 (1.10)</td>
<td>0.0001</td>
</tr>
<tr>
<td>HDL cholesterol, mmol/L</td>
<td>1.52 (1.28)</td>
<td>1.50 (1.28)</td>
<td>1.51 (1.28)</td>
<td>1.53 (1.28)</td>
<td>0.009</td>
</tr>
<tr>
<td>Triglycerides, mmol/L</td>
<td>1.72 (1.85)</td>
<td>1.72 (1.81)</td>
<td>1.68 (1.81)</td>
<td>1.66 (1.82)</td>
<td>0.06</td>
</tr>
<tr>
<td>HbA1c, % total</td>
<td>5.24 (1.11)</td>
<td>5.23 (1.11)</td>
<td>5.23 (1.12)</td>
<td>5.20 (1.10)</td>
<td>0.08</td>
</tr>
<tr>
<td>Fibrinogen, g/L</td>
<td>2.89 (1.23)</td>
<td>2.94 (1.23)</td>
<td>2.90 (1.22)</td>
<td>2.87 (1.22)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>CRP, mg/L</td>
<td>1.03 (0.27)</td>
<td>1.10 (0.31)</td>
<td>1.08 (0.36)</td>
<td>0.96 (0.39)</td>
<td>0.0001</td>
</tr>
<tr>
<td>D-dimer, ng/mL</td>
<td>157.7 (1.8)</td>
<td>166.3 (1.8)</td>
<td>161.5 (1.8)</td>
<td>160.0 (1.8)</td>
<td>0.02</td>
</tr>
<tr>
<td>t-PA, ng/mL</td>
<td>4.54 (1.75)</td>
<td>4.64 (1.73)</td>
<td>4.57 (1.75)</td>
<td>4.38 (1.74)</td>
<td>0.0003</td>
</tr>
<tr>
<td>vWF, IU/dL</td>
<td>115.9 (1.4)</td>
<td>118.7 (1.4)</td>
<td>115.6 (1.4)</td>
<td>114.3 (1.4)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Prevalence by infant feeding status, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity (BMI $\geq$ 30 kg/m$^2$)</td>
<td>25</td>
<td>28</td>
<td>24</td>
<td>22</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

a Breastfeeding status unknown either because it was not known at age 7 or because the cohort member joined the study at a later date.

b $P$ for trend is across 3 categories of known infant feeding.

c GSD is the exponentiated SD of the log-transformed values, the 95% range for the geometric mean is from (geometric mean $\times$ GSD$^{1/2}$) to (geometric mean $\times$ GSD$^{1/2}$).
tion of breastfeeding showed a weak inverse association with hip circumference \((P\text{ for trend } = 0.06)\), triglycerides \((P = 0.06)\), and HbA1c \((P = 0.08)\). Infant feeding status was not associated with blood pressure (systolic or diastolic), total or LDL cholesterol, or height-adjusted FEV\(_1\) (Table 1).

There was no difference in the proportion of men (46%) or women (46%) breastfed for more than 1 month. Breastfeeding was more common in higher social classes (similar patterns were observed by childhood and adulthood social economic status) and in the southern part of England compared with central regions, Wales, or Scotland (Table 2). Mothers who did not smoke during pregnancy were more likely to breastfeed for \(>1\) month (50% compared with 40% of medium smokers and 37% of heavy smokers among those with data on infant feeding). Mothers with relatively lower prepregnancy body size were also more likely to breastfeed. Subjects with higher birth weight were more likely to have been breastfed. Adult current smokers were less likely to have been breastfed for \(>1\) month compared with never-smokers (42% versus 46% among those with data on infant feeding).

After adjustment for the confounders mentioned above, the apparent beneficial effect of breastfeeding over formula feeding on cardiorespiratory risk factors was attenuated (Table 3). However, being breastfed for \(>1\) month, compared with formula feeding, was associated with decreased waist circumference (\(-0.81\%\); 95% confidence interval [CI]: \(-1.47\%\) to \(-0.14\%\)), waist/hip ratio (\(-0.54\%\); 95% CI: \(-0.92\%\) to \(-0.16\%\)), reduced risk of obesity (relative risk: 0.85; 95% CI: 0.75 to 0.97), and lower levels of fibrinogen (\(-1.43\%\); 95% CI: \(-2.55\%\) to \(-0.30\%\)), vWF (\(-2.15\%\); 95% CI: \(-4.01\%\) to \(-0.26\%\)), and marginally lower t-PA (\(-2.75\%\); 95% CI: \(-5.54\%\) to 0.11%), all showing statistically significant inverse trends across known feeding groups (\(P\)s from 0.004 to 0.05).

There was no evidence of an interaction between breastfeeding and gender for any risk factor except for fibrinogen and CRP (interaction \(P = 0.05\) and .02, respectively). The association of fibrinogen and CRP with breastfeeding seemed to be restricted to women, with no appreciable effect in men. In women, after adjustment, fibrinogen levels were 2.6% lower (95% CI: 4.6% to 1.0% lower; \(P = .003\)) in those breastfed for \(>1\) month compared with those not breastfed, and CRP levels were 12.8% lower (95% CI: 21.5% to 3.0% lower; \(P = .01\)).

To examine whether the inverse association between breastfeeding and waist circumference, waist/hip ratio, obesity, and the inflammatory makers might be explained by residual confounding, a number of additional factors related to childhood social and cultural environment were identified and adjusted for, including maternal age, level of paternal education, and number of people per household. Adult smoking status was further refined by taking account of the number of cigarettes smoked. Inclusion of these factors in the regression model did not materially alter the findings (data not presented). Current smoking is a strong predictor of the levels of fibrinogen, CRP, and vWF, with smokers having higher levels than former- or never-smokers. Restricting the analyses to former- or never-smokers did not markedly change the results; in women, fibrinogen levels were 1.9% lower and CRP levels 13.6% lower with breastfeeding for \(>1\) month compared with formula feeding. For both men and women, vWF was 2.0% lower and 1.3% lower with breastfeeding for \(<1\) month and \(>1\) month, respectively, compared with formula feeding.

DISCUSSION

In our study, individuals breastfed for \(>1\) month were, on average, taller in adulthood; had lower BMI, waist and hip circumference, waist/hip ratio, total and LDL cholesterol, triglycerides, HbA1c, fibrinogen, t-PA, vWF; were less likely to be obese; and higher blood pressure, FEV\(_1\), and HDL cholesterol compared with those who were not breastfed. Those breastfed for \(<1\) month tended to have levels between these 2 groups. The inverse association of breastfeeding duration on established cardiovascular risk factors was attenuated after adjustment for confounding factors. However, the associations with waist circumference (0.8% lower), waist/hip ratio (0.5% lower), obesity (15% lower odds), fibrinogen (3% lower in women), CRP (13% lower in women), t-PA (3% lower), and vWF (2% lower) remained statistically significant after adjustment.

The graded effect across the 3 feeding groups on adult height was markedly reduced after adjustment. This agrees with previous work\(^\text{10,18,22}\) that has shown no association between infant feeding and attained height in adulthood. A recent meta-analysis found a small effect of breastfeeding on BMI, but this effect was attenuated to the null after adjustment for potentially important confounders (socioeconomic status, maternal smoking, and BMI in early life).\(^\text{19}\) Similarly, findings from the current study found no association between infant feeding and adult BMI after adjustment. Reports of an association between breastfeeding and obesity in adulthood vary.\(^\text{10,13,25,26}\) We found the odds of obesity (individuals with BMI \(\geq 30\text{ kg/m}^2\)) to be 15% lower in those breastfed for \(>1\) month compared with those formula fed (adjusted odds ratio: 0.85; 95% CI: 0.75 to 0.97). Our result is similar to findings reported on the same cohort earlier in adulthood\(^\text{19}\) and agrees with a systematic review reporting a small effect of breastfeeding on the odds of obesity after adjustment (adjusted odds ratio: 0.93; 95% CI: 0.88 to 0.99).\(^\text{17}\) The prevalence of exclusive formula feeding is approximately one third in Europe and the United States and agrees with the prevalence of not being breastfed in the current study (30%). Therefore, a 15% reduction in obesity could equate to approx-
<table>
<thead>
<tr>
<th>Variable, n (%)</th>
<th>Not Known&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Not Breastfed</th>
<th>Breastfed &lt;1 mo</th>
<th>Breastfed &gt;1 mo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender&lt;sup&gt;a&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>624 (52)</td>
<td>1229 (50)</td>
<td>945 (48)</td>
<td>1867 (50)</td>
</tr>
<tr>
<td>Female</td>
<td>581 (48)</td>
<td>1211 (50)</td>
<td>1015 (52)</td>
<td>1905 (50)</td>
</tr>
<tr>
<td><strong>Region of birth&lt;sup&gt;a&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scotland</td>
<td>97 (8.1)</td>
<td>400 (16)</td>
<td>154 (7.9)</td>
<td>280 (7.4)</td>
</tr>
<tr>
<td>Northern</td>
<td>37 (3.1)</td>
<td>240 (9.8)</td>
<td>162 (8.3)</td>
<td>216 (5.7)</td>
</tr>
<tr>
<td>East and West Ridings</td>
<td>85 (7.0)</td>
<td>207 (8.5)</td>
<td>177 (9.0)</td>
<td>301 (8.0)</td>
</tr>
<tr>
<td>Northwestern</td>
<td>112 (9.3)</td>
<td>286 (12)</td>
<td>279 (14)</td>
<td>453 (12)</td>
</tr>
<tr>
<td>North Midlands</td>
<td>72 (6.0)</td>
<td>193 (7.9)</td>
<td>158 (8.1)</td>
<td>264 (7.0)</td>
</tr>
<tr>
<td>Wales</td>
<td>54 (4.5)</td>
<td>182 (7.5)</td>
<td>107 (5.5)</td>
<td>140 (3.7)</td>
</tr>
<tr>
<td>Midlands</td>
<td>84 (7.0)</td>
<td>209 (8.6)</td>
<td>215 (11)</td>
<td>338 (9.0)</td>
</tr>
<tr>
<td>Eastern</td>
<td>54 (4.5)</td>
<td>159 (6.5)</td>
<td>138 (7.0)</td>
<td>344 (9.1)</td>
</tr>
<tr>
<td>Southwestern</td>
<td>65 (5.4)</td>
<td>124 (5.1)</td>
<td>97 (5.0)</td>
<td>235 (6.2)</td>
</tr>
<tr>
<td>Southern</td>
<td>67 (5.6)</td>
<td>104 (4.3)</td>
<td>100 (5.1)</td>
<td>276 (7.3)</td>
</tr>
<tr>
<td>South East</td>
<td>42 (3.5)</td>
<td>91 (3.7)</td>
<td>90 (4.6)</td>
<td>251 (6.6)</td>
</tr>
<tr>
<td>Greater London</td>
<td>197 (16)</td>
<td>221 (9.1)</td>
<td>259 (13.2)</td>
<td>607 (16.1)</td>
</tr>
<tr>
<td>Unknown</td>
<td>239 (20)</td>
<td>24 (1.0)</td>
<td>24 (1.2)</td>
<td>67 (1.8)</td>
</tr>
<tr>
<td><strong>Socioeconomic position in childhood&lt;sup&gt;a&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional</td>
<td>69 (5.7)</td>
<td>91 (3.7)</td>
<td>76 (3.9)</td>
<td>242 (6.4)</td>
</tr>
<tr>
<td>Managerial, technical</td>
<td>116 (9.6)</td>
<td>311 (13)</td>
<td>233 (12)</td>
<td>634 (17)</td>
</tr>
<tr>
<td>Skilled nonmanual</td>
<td>90 (7.5)</td>
<td>210 (8.6)</td>
<td>169 (8.6)</td>
<td>442 (12)</td>
</tr>
<tr>
<td>Skilled manual</td>
<td>433 (36)</td>
<td>1230 (50)</td>
<td>1028 (52)</td>
<td>1790 (47)</td>
</tr>
<tr>
<td>Partly skilled</td>
<td>114 (9.5)</td>
<td>351 (14)</td>
<td>259 (13)</td>
<td>404 (11)</td>
</tr>
<tr>
<td>Unskilled</td>
<td>80 (6.6)</td>
<td>224 (9.2)</td>
<td>184 (9.4)</td>
<td>246 (6.5)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (0.2)</td>
<td>16 (0.7)</td>
<td>9 (0.5)</td>
<td>13 (0.3)</td>
</tr>
<tr>
<td>Unknown</td>
<td>301 (25)</td>
<td>7 (0.3)</td>
<td>2 (0.1)</td>
<td>1 (0.03)</td>
</tr>
<tr>
<td><strong>Maternal smoking during pregnancy&lt;sup&gt;a&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonsmoker</td>
<td>603 (50)</td>
<td>1469 (60)</td>
<td>1221 (62)</td>
<td>2650 (70)</td>
</tr>
<tr>
<td>Medium</td>
<td>159 (13)</td>
<td>380 (16)</td>
<td>311 (16)</td>
<td>450 (12)</td>
</tr>
<tr>
<td>Heavy</td>
<td>119 (9.9)</td>
<td>332 (14)</td>
<td>223 (11)</td>
<td>331 (8.8)</td>
</tr>
<tr>
<td>Variable</td>
<td>50 (4.2)</td>
<td>155 (6.4)</td>
<td>126 (6.4)</td>
<td>193 (5.1)</td>
</tr>
<tr>
<td>Unknown</td>
<td>274 (23)</td>
<td>104 (4.3)</td>
<td>79 (4.0)</td>
<td>148 (3.9)</td>
</tr>
<tr>
<td><strong>Maternal weight, kg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;51</td>
<td>158 (13)</td>
<td>362 (15)</td>
<td>310 (16)</td>
<td>463 (12)</td>
</tr>
<tr>
<td>51 &lt; 64</td>
<td>537 (4.5)</td>
<td>1246 (51)</td>
<td>1054 (54)</td>
<td>2139 (57)</td>
</tr>
<tr>
<td>64 &lt; 89</td>
<td>211 (18)</td>
<td>662 (27)</td>
<td>477 (24)</td>
<td>968 (26)</td>
</tr>
<tr>
<td>&gt;89</td>
<td>12 (1.0)</td>
<td>39 (1.6)</td>
<td>23 (1.2)</td>
<td>30 (1.0)</td>
</tr>
<tr>
<td>Unknown</td>
<td>287 (24)</td>
<td>131 (5.4)</td>
<td>96 (4.9)</td>
<td>172 (4.6)</td>
</tr>
<tr>
<td><strong>Birth weight, kg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;3</td>
<td>257 (21)</td>
<td>642 (26)</td>
<td>442 (23)</td>
<td>728 (19)</td>
</tr>
<tr>
<td>3 &lt; 3.3</td>
<td>271 (18)</td>
<td>516 (21)</td>
<td>492 (25)</td>
<td>860 (23)</td>
</tr>
<tr>
<td>3.3 &lt; 3.6</td>
<td>177 (15)</td>
<td>397 (16)</td>
<td>354 (18)</td>
<td>785 (21)</td>
</tr>
<tr>
<td>3.6–3.9</td>
<td>148 (12)</td>
<td>433 (18)</td>
<td>352 (18)</td>
<td>678 (18)</td>
</tr>
<tr>
<td>&gt;3.9</td>
<td>117 (9.7)</td>
<td>290 (12)</td>
<td>214 (11)</td>
<td>495 (13)</td>
</tr>
<tr>
<td>Unknown</td>
<td>289 (24)</td>
<td>162 (6.6)</td>
<td>106 (5.4)</td>
<td>226 (6.0)</td>
</tr>
<tr>
<td><strong>Birth weight, mean (SD), kg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional</td>
<td>3.30 (0.53)</td>
<td>3.30 (0.57)</td>
<td>3.33 (0.49)</td>
<td>3.38 (0.48)</td>
</tr>
<tr>
<td>Managerial, technical</td>
<td>77 (6.4)</td>
<td>91 (3.7)</td>
<td>81 (4.1)</td>
<td>243 (6.4)</td>
</tr>
<tr>
<td>Skilled nonmanual</td>
<td>417 (33)</td>
<td>791 (32)</td>
<td>643 (33)</td>
<td>1433 (38)</td>
</tr>
<tr>
<td>Skilled manual</td>
<td>235 (20)</td>
<td>500 (20)</td>
<td>425 (22)</td>
<td>775 (21)</td>
</tr>
<tr>
<td>Partly skilled</td>
<td>249 (21)</td>
<td>490 (20)</td>
<td>387 (20)</td>
<td>617 (16)</td>
</tr>
<tr>
<td>Unskilled</td>
<td>130 (11)</td>
<td>340 (14)</td>
<td>262 (13)</td>
<td>428 (11)</td>
</tr>
<tr>
<td>Other</td>
<td>43 (3.6)</td>
<td>101 (4.1)</td>
<td>75 (3.8)</td>
<td>109 (2.9)</td>
</tr>
<tr>
<td>Unknown</td>
<td>50 (4.2)</td>
<td>124 (5.1)</td>
<td>76 (3.9)</td>
<td>158 (4.2)</td>
</tr>
<tr>
<td><strong>Smoking status in adulthood&lt;sup&gt;a&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>519 (43)</td>
<td>1095 (45)</td>
<td>869 (44)</td>
<td>1673 (44)</td>
</tr>
<tr>
<td>Former/occasional</td>
<td>354 (29)</td>
<td>643 (26)</td>
<td>544 (28)</td>
<td>1180 (31)</td>
</tr>
<tr>
<td>Current</td>
<td>300 (25)</td>
<td>616 (25)</td>
<td>490 (23)</td>
<td>796 (21)</td>
</tr>
<tr>
<td>Unknown</td>
<td>32 (2.7)</td>
<td>86 (3.5)</td>
<td>57 (2.9)</td>
<td>123 (3.3)</td>
</tr>
<tr>
<td><strong>Total, N</strong></td>
<td>1205</td>
<td>2440</td>
<td>1960</td>
<td>3772</td>
</tr>
</tbody>
</table>

<sup>a</sup> Percentages >10% are rounded to nearest whole percent.

<sup>b</sup> Breastfeeding status unknown either because it was not known at age 7 or because the cohort member joined the study at a later date.
imately a 4% reduction in the burden of obesity if all mothers were to breastfeed. Based on observational evidence, the relative risk of coronary heart disease27–30 or cardiovascular disease31 outcomes in obese versus non-obese individuals is approximately double the risk of type 2 diabetes.39 These associations are manifest by approximately a 10% to 15% variation in waist circumference or waist/hip ratio. In the current study, those breastfed for >1 month had reduced waist circumference or waist/hip ratio by <1%. Therefore, the potential protective effect of breastfeeding for >1 month on cardiovascular disease and diabetes via any effect on waist circumference or waist/hip ratio is vanishingly small.

The absence of an association with blood pressure is in agreement with 2 systematic reviews36,40 that found no effect in larger studies (with ≥1000 participants). Evidence of small study bias was observed in these re-

---

**TABLE 3** Comparison of Cardiorespiratory Risk Factors in Adults Who Were Breastfed for >1 Month or <1 Month and Those Never Breastfed

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Adjusted Difference (95% CI)</td>
<td>Adjusted Difference (95% CI)</td>
<td></td>
</tr>
<tr>
<td>Height, cm</td>
<td>8045</td>
<td>0.11 (−0.26 to 0.47)</td>
<td>0.18 (−0.14 to 0.50)</td>
<td>.28</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>8032</td>
<td>−0.12 (−0.42 to 0.18)</td>
<td>−0.22 (−0.48 to 0.04)</td>
<td>.10</td>
</tr>
<tr>
<td>Systolic blood pressure, mm Hg</td>
<td>7008</td>
<td>−0.90 (−1.87 to 0.07)</td>
<td>0.30 (−0.55 to 1.16)</td>
<td>.34</td>
</tr>
<tr>
<td>Diastolic blood pressure, mm Hg</td>
<td>7008</td>
<td>−0.42 (−1.08 to 0.24)</td>
<td>0.15 (−0.43 to 0.73)</td>
<td>.50</td>
</tr>
<tr>
<td>Total cholesterol, mmol/L</td>
<td>6723</td>
<td>0.01 (−0.07 to 0.08)</td>
<td>−0.01 (−0.08 to 0.05)</td>
<td>.60</td>
</tr>
<tr>
<td>LDL cholesterol, mmol/L</td>
<td>6361</td>
<td>0.00 (−0.06 to 0.06)</td>
<td>−0.02 (−0.08 to 0.03)</td>
<td>.37</td>
</tr>
<tr>
<td>Height adjusted FEV₁, L</td>
<td>7314</td>
<td>0.02 (−0.01 to 0.05)</td>
<td>0.02 (−0.01 to 0.05)</td>
<td>.15</td>
</tr>
</tbody>
</table>

Adjusted mean difference

- Height, cm: 0.11 (−0.26 to 0.47)
- BMI, kg/m²: −0.12 (−0.42 to 0.18)
- Systolic blood pressure, mm Hg: −0.90 (−1.87 to 0.07)
- Diastolic blood pressure, mm Hg: −0.42 (−1.08 to 0.24)
- Total cholesterol, mmol/L: 0.01 (−0.07 to 0.08)
- LDL cholesterol, mmol/L: 0.00 (−0.06 to 0.06)
- Height adjusted FEV₁, L: 0.02 (−0.01 to 0.05)

Adjusted odds ratio (95% CI)

| Waist circumference, cm | 7808 | −0.20 (−0.96 to 0.01) | −0.81 (−1.47 to −0.14) | .01 |
| Hip circumference, cm   | 7769 | −0.13 (−0.65 to 0.39) | −0.31 (−0.77 to 0.14)  | .17 |
| Waist/hip ratio         | 7667 | −0.09 (−0.52 to 0.35) | −0.54 (−0.92 to 0.16)  | .004 |
| HDL cholesterol, mmol/L | 6710 | 0.29 (−1.21 to 1.81)  | 0.56 (−0.76 to 1.90)   | .40 |
| Triglycerides, mmol/L   | 6702 | −0.53 (−4.02 to 3.08) | −0.15 (−3.22 to 3.02) | .94 |
| HbA1c, % total          | 6818 | 0.31 (−0.37 to 1.00)  | 0.05 (−0.54 to 0.60)   | .93 |
| Fibrinogen, g/L         | 6611 | −0.96 (−2.23 to 0.33) | −1.43 (−2.55 to −0.30) | .01 |
| D-dimer, mg/mL          | 6584 | −3.16 (−6.60 to 0.41) | −2.81 (−5.85 to 0.32)  | .10 |
| vWF, IU/dL              | 6620 | −2.52 (−4.61 to −0.38) | −2.15 (−4.01 to −0.26) | .04 |

**Footnotes:**

- Data are adjusted for social economic position in childhood and adulthood, gender, region of birth, smoking in adulthood, birth weight, pre-pregnancy maternal body size, and maternal smoking during pregnancy, plus the following: height; nurse; BMI/obesity: weight measured on hard floor/carpet, recent food consumption, nurse; BP: medication for blood pressure reduction, recent food consumption, ambient air temperature, time of day, month of examination, nurse, and instrument; total and LDL cholesterol: delay in processing blood sample and month of examination; FEV₁: height, use of inhalers within the last 24 hours, upper respiratory tract infection in the last 3 weeks, month of examination, and instrument; waist circumference: nurse; hip circumference: nurse; waist/hip ratio: nurse; HDL cholesterol: delay in processing blood sample, recent food consumption, time of day, and month of examination; triglycerides: delay in processing blood sample, recent food consumption, and time of day; HbA1c: treatment for diabetes, month of examination, and sample delay; fibrinogen, CRP, and D-dimer: time of day, month of examination, and laboratory batch; vPA and vWF: delay in processing blood sample, time of day, month of examination, and laboratory batch.

- Adjusted mean difference: Average difference in the measure of interest between those breastfed and those formula fed, adjusted for confounding factors.

- Adjusted odds ratio: Ratio of odds of a cardiovascular or metabolic outcome in those breastfed compared to those formula fed, adjusted for confounding factors.
views, indicating the presence of publication bias. The authors concluded that the pooled association between breastfeeding and systolic blood pressure may have been exaggerated by large effects reported in smaller studies. Although this study excludes any important lowering effect of breastfeeding on blood pressure, it can not rule out that prolonged exclusive breastfeeding may have a protective effect.41 Studies with follow-up in childhood and adolescence have shown a dose-response relationship between duration of exclusive breastfeeding (rather than mixed feeding) and systolic blood pressure;24,43; whether this effect persists into adulthood remains unclear.25 In the current study it was not possible to examine exclusivity of breastfeeding or duration of breastfeeding beyond 1 month.

The lack of an association between lipids and breastfeeding after adjustment disagrees with a meta-analysis that found total and LDL cholesterol levels to be ~0.2 mmol/L lower in those breastfed.5 However, it has been suggested that the beneficial effect on lipids requires exclusive breastfeeding for >3 months,44 and it was not possible to examine this in the current study. In agreement with a previous study, we did not find any beneficial effect of breastfeeding on lung function.15

The graded inverse association between breastfeeding and levels of inflammatory markers in the current study was attenuated by adjustment for potential confounders but remained statistically significant for fibrinogen and vWF and marginally significant for t-PA. The differential effect of breastfeeding on fibrinogen and CRP between men and women, with stronger inverse associations in women, was unexpected and may be a chance finding (interaction \( P = .05 \) and .02, respectively). The only other study examining fibrinogen levels in adulthood and breastfeeding did not find an association,25 but infant feeding was recalled in adulthood (some 50 years after birth), and the authors acknowledged that selection and recall bias may have affected their results. A long-term follow-up of preterm births that were randomly assigned to breast milk or formula feed exhibited lower levels of CRP in adolescence,45 and lower levels of CRP were found in women (but not men) breastfed for longer durations.21 A lower level of vWF in both men and women in those breast compared with formula feeding is a novel finding. The magnitude of the associations between duration of breastfeeding and levels of the inflammatory markers was not materially altered after exclusion of current smokers, but any potential influence of infant feeding on coronary heart disease incidence is likely to be small.46–49

Prolonged and exclusive breastfeeding has been associated with improved neurocognitive development.50–55 Because socioeconomic status is related to both cognition and health status, adjustment for markers of adult socioeconomic position may not be appropriate, because this may lie on the causal pathway. Although it was only the association of infant feeding with BMI, waist circumference, and waist/hip ratio that were marginally strengthened without adjustment for socioeconomic status, the potential effects of mode of early feeding on cardiovascular27,29,31,37 or diabetic39 outcomes are still modest (<3%) and remain of limited public health importance.

A potential weakness of the current study was that breastfeeding status was not reported as being exclusive and was recalled after 7 years by parental interview. Approximately half of the mothers (46%) reported breastfeeding wholly or partly for >1 month in this study, which agrees well with a report in a specific region of England at this time, where 54% had breastfed for \( \geq 1 \) month.56 In terms of the validity of recalling infant feeding status, this has been shown to be accurate \( \leq 20 \) years after birth.57 Although no information on the exclusivity of breastfeeding in the United Kingdom at this time could be found, formula feeding in the late 1950s and early 1960 would probably have consisted of dried cow’s milk fortified with iron and vitamins.58 Although the content of formula milk has evolved since this time,59 we believe that our findings would not be materially changed by the introduction of modern formulae in more contemporary cohorts. The British 1958 birth cohort is predominantly white (97%); hence, it was not possible to examine racial differences in the association between infant feeding patterns and cardiorespiratory risk in adult life.

CONCLUSIONS

Breastfeeding is beneficial for many other reasons, but breastfeeding for >1 month is unlikely to offer substantial protection against cardiorespiratory disease in adult life. The association between breastfeeding and waist circumference, waist/hip ratio, and obesity is of interest and needs to be replicated by other studies that have information on exclusive breastfeeding for longer durations (>1 month). Breastfeeding is associated with early life social environment, which influences adult behavior and lifestyle, and it is important that such confounders operating across the life course are identified and taken into consideration when examining the association between infant feeding and cardiorespiratory outcomes in later life. Although the potential effect of mode of early feeding on inflammatory markers maybe of limited relevance to cardiovascular outcomes, it may be of relevance to other disease mechanisms.

ACKNOWLEDGMENTS

The medical examination in adulthood and related statistical analyses were funded by Medical Research Council grant G0000934, awarded under the Health of the Public initiative. The Medical Research Council played no role in study design, collection, analysis and interpretation of data, writing the report, or submitting the
REFERENCES


years of follow-up of the participants in the study of men born
52. Horwood LJ, Fergusson DM. Breastfeeding and later cognitive and academic outcomes. Pediatrics. 1998;101(1). Available at: www.pediatrics.org/cgi/content/full/101/1/e9
Long-term Follow-up of 414 HIV-Infected Romanian Children and Adolescents Receiving Lopinavir/Ritonavir-Containing Highly Active Antiretroviral Therapy

Mark W. Kline, MDa,b, Sorin Rugina, MDc, Margareta Ilie, MDc, Rodica F. Matusa, MDb, Ana-Maria Schweitzer, MAa,b, Nancy R. Calles, BSN, RNa,b, Heidi L. Schwarzwald, MDa,b

aBaylor College of Medicine International Pediatric AIDS Initiative, Texas Children's Hospital, Houston, Texas; bBaylor Black Sea Foundation, Constanta, Romania; cInfectious Diseases Hospital, Constanta, Romania

ABSTRACT

BACKGROUND. There are no published reports of the long-term safety and effectiveness of highly active antiretroviral therapy for children and adolescents living in resource-limited settings or of large cohorts of HIV-infected children and adolescents treated long-term (>48 weeks) with lopinavir/ritonavir-containing highly active antiretroviral therapy.

OBJECTIVES. The purpose of this work was to evaluate the long-term outcomes of treatment of HIV-infected children and adolescents with lopinavir/ritonavir-containing highly active antiretroviral therapy in a resource-limited setting.

METHODS. We studied an inception cohort of 414 HIV-infected children receiving lopinavir/ritonavir-containing highly active antiretroviral therapy between November 2001 and August 2006 at the Romanian-American Children’s Center in Constanta, Romania. The center provides comprehensive primary and HIV specialty care and treatment to all known HIV-infected children and adolescents living in Constanta. We measured safety and effectiveness by the percentage of children remaining on treatment, rates of mortality, and changes in plasma HIV RNA concentrations and CD4+ lymphocyte counts.

RESULTS. The study population consisted predominantly of antiretroviral drug–experienced older children and adolescents with advanced HIV disease. Treatment was well tolerated, with 337 children (81%) remaining on therapy after a median duration of >4 years. Thirty-seven deaths occurred; the death rate compared favorably to prospectively collected historical data. The most recent on-treatment plasma HIV RNA concentration was <400 copies per milliliter in 192 of 265 children tested. The mean baseline CD4+ lymphocyte count was 292 cells per microliter (n = 299); the mean change from baseline was +266 (n = 284), +317

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2802
doi:10.1542/peds.2006-2802

Key Words
HIV infection, child, adolescent, lopinavir

Abbreviations
HAART — highly active antiretroviral therapy
BIPAI — Baylor College of Medicine International Pediatric AIDS Initiative
Accepted for publication Nov 13, 2006
Address correspondence to Mark W. Kline, MD, Department of Pediatrics, Baylor College of Medicine, 6621 Fannin St, CCC1210, Houston, TX 77030. E-mail: mkline@bcm.edu
PEDiATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
More than 90% of these children live in low- and middle-income countries. There are no published reports of the long-term safety and effectiveness of highly active antiretroviral therapy (HAART) for children in these settings. Lopinavir/ritonavir-containing highly active antiretroviral therapy is a safe, effective, and durable treatment option for antiretroviral drug–experienced older children and adolescents with advanced HIV disease.

The World Health Organization recently estimated that 800,000 children <15 years of age are in immediate need of antiretroviral treatment. More than 90% of these children live in low- and middle-income countries. There are no published reports of the long-term safety and effectiveness of highly active antiretroviral therapy (HAART) for children in these settings.

Lopinavir/ritonavir (Kaletra)-containing HAART is strongly recommended for initial therapy of HIV infection in children by the Working Group on Antiretroviral Therapy and Medical Management of HIV-Infected Children of the US Health Resources and Services Administration and National Institutes of Health. Approval of the drug for children was based largely on the results of a 48-week study of its safety and efficacy in 100 HIV-infected children 6 months to 12 years of age. There are no published reports of the long-term (>48-week) safety and effectiveness of lopinavir/ritonavir-containing HAART in large cohorts of HIV-infected children and adolescents.

Romania is a middle-income country with a per capita gross domestic product comparable to that of Botswana or Thailand and total health expenditure per capita less than Botswana, Thailand, or South Africa. Thousands of Romanian children were infected horizontally with HIV in the late 1980s through the transfusion of whole human blood unscreened for HIV, often by reuse of disposable needles.

The Baylor College of Medicine International Pediatric AIDS Initiative (BIPAI), Texas Children’s Hospital, the Infectious Diseases Hospital Constanta, and the Romanian Ministry of Health and Family built and opened the Infectious Diseases Hospital-BIPAI Outpatient HIV Clinic (also known as the Romanian-American Children’s Center) in Constanta in April 2001. In a unique partnership arrangement, the government of Romania supports the medical care and antiretroviral treatment of HIV-infected children and adolescents in the center through its national program, complemented by BIPAI and Baylor Black Sea Foundation–supported comprehensive medical and psychosocial services and antiretroviral drug donations.

A program of child and adolescent HAART was initiated at the Romanian-American Children’s Center in November 2001. Today, 476 children and adolescents receive HAART at the center, more than in any other European center. We report the long-term outcomes of treatment in 414 HIV-infected children and adolescents who received lopinavir/ritonavir-containing HAART at the Romanian-American Children’s Center between November 2001 and August 2006.

METHODS

Criteria for Treatment and Drug Administration
Children and adolescents received lopinavir/ritonavir-containing HAART in an open manner, by prescription, according to package insert instructions. Children in any US Centers for Disease Control and Prevention category other than N1 or A1 were considered eligible for treatment. Children already receiving antiretroviral therapy were considered eligible for a change to lopinavir/ritonavir-containing HAART if they had evidence of clinical, virologic, and immunologic HIV disease progression.

Drugs used in combination with lopinavir/ritonavir were chosen on the basis of patient treatment history; genotyping for resistance mutations was not performed. Every child was naïve to lopinavir/ritonavir and to ≥1 other agent that was being used in combination. Many children had been treated previously with other HIV protease inhibitors.

All of the antiretroviral medications were stored according to package insert instructions. Weight and body surface area were determined every 4 weeks, and the drug dose was changed when indicated by an increase or decrease in weight. Adherence was monitored by measuring or counting returned medication and by questioning the parent and/or child at each scheduled clinic visit.

Patients received prophylaxis for Pneumocystis jiroveci pneumonia according to established guidelines, and nutritional support and antibiotic therapy were prescribed as needed. Immunomodulators (excluding corticosteroids and intravenous immunoglobulin) were not used.

Patient and Family Education and Support
All of the patients and caregivers were counseled initially and monthly regarding the importance of medication adherence. At the time that HAART was initiated, each child and family was given a 1-page, individualized reference sheet showing actual-size, color photographs of the pills that they would be taking, together with dosing instructions and any other special instructions (eg, diet). Medication boxes were provided, and each family was instructed on their correct use. An illustrated 20-page Romanian language booklet on HIV and antiretroviral therapy was created for use at the center, and it was given to each family as an adjunct to verbal
counseling. This booklet contains a glossary of key terms and is written at a sixth- or seventh-grade educational level. A mobile, multidisciplinary medical/psychosocial team visited children and families in the community ~5 days each week. Center staff identified children and families who they felt might benefit from this more intensive schedule of follow-up at home and in the community. These home visits permitted monitoring of antiretroviral medication storage, administration and adherence, and factors in the home and family that might impact treatment. Five separate parent-led support groups met on a regular basis with members of the health care team. These meetings provided a forum for discussion of a variety of medical, psychosocial, and other issues and served as an immense source of support for families coping with the care of chronically ill children and adolescents.

Clinical and Laboratory Monitoring
All of the treated children and adolescents were evaluated clinically and with routine laboratory tests on a monthly basis. The Division of AIDS Toxicity Table for Grading Severity of Pediatric Adverse Experiences was used as a rough guide for the management of presumed drug-associated toxicities. In brief, this table grades a variety of potential study drug-associated clinical and laboratory adverse events on a 4-point scale from grade 1 (least severe) to grade 4 (most severe). Laboratory values constituting grade ≥3 abnormalities include the following: hemoglobin concentration, <7 g/dL; absolute neutrophil count, <400/μL; alanine aminotransferase, ≥10 times the upper limit of normal; bilirubin, ≥3 times the upper limit of normal; and serum creatinine, >1.1 mg/dL (age: ≥3 months to 2 years) or 1.6 mg/dL (age: ≥2 years). Grade ≥3 instances of presumed study drug-associated toxicity or intolerance are managed according to a dose-modification scheme that mandates the interruption of antiretroviral drug therapy for ≥28 days. If the toxicity improves during that time to less than grade 3, therapy is resumed.

Because of cost considerations, plasma HIV RNA measurements (Roche Molecular Systems, Branchburg, NJ) were performed only on a subset of treated children and adolescents until ~2005, when the test became available more routinely. In general, CD4+ lymphocyte counts were obtained at baseline and intervals of ~6 to 12 months.

RESULTS
Study Population
A total of 414 children and adolescents received lopinavir/ritonavir-containing HAART between November 2001 and August 2006. Table 1 shows selected characteristics of lopinavir/ritonavir-treated children at baseline. Table 2 shows the drugs used in combination with lopinavir/ritonavir. The median duration of follow-up was 51 months (range: 8–57 months). Table 3 shows the current treatment status of these children and adolescents.

Safety and Clinical Observations
Lopinavir/ritonavir-containing HAART was well tolerated, with 337 (81%) of 414 children remaining on therapy after a median duration of >4 years. Treatment with lopinavir/ritonavir was interrupted on 312 occasions in 220 children. The most common reasons for treatment interruption were clinical adverse events or laboratory abnormalities (87 episodes), temporary unavailability of concomitantly administered antiretroviral medications (63 episodes), death (37 episodes), HIV disease progression (25 episodes), and medication nonadherence (23 episodes). Few clinical adverse events or laboratory abnormalities of at least moderate severity (grade ≥3) and of possible or probable relationship to lopinavir/ritonavir therapy were observed, consisting principally of intractable vomiting (11 episodes) and hepatitis (3 episodes). Treatment was interrupted in 11 cases because of adverse events or intolerance attributed to concomitantly administered antiretroviral drugs. Six cases of lipodystrophy were diagnosed. Seventy-seven children permanently discontinued therapy with lopinavir/ritonavir (Table 3).

Thirty-seven deaths occurred among children receiv-
ing lopinavir/ritonavir-containing HAART. Causes of death included tuberculosis (8 cases), progressive encephalopathy (4 cases), cytomegalovirus disease (3 cases), and pneumonia (2 cases). A specific cause of death was not known in approximately half of the cases; none of the deaths was thought to be related in any way to antiretroviral therapy. Most of the deaths occurred among children with advanced HIV disease at the time that HAART was initiated; 14 deaths (38%) occurred within 16 weeks of starting HAART. Twenty-five deaths occurred during the first half of the study period (November 2001 through December 2003); only 12 deaths occurred during the second half of the study period (January 2004 through August 2006). The mortality rate observed during these 2 time periods was ~3.3 per 100 patient-years and 1.2 per 100 patient-years, respectively. We observed an average annual mortality rate of ~13% in a cohort of HIV-infected children followed prospectively in Constanta over a 4-year period ending in May 2002.7

Virologic and Immunologic Observations

Only 97 children and adolescents had plasma HIV RNA concentrations measured immediately before beginning lopinavir/ritonavir-containing HAART. In this subgroup, the mean baseline plasma HIV RNA concentration was 152,036 copies per milliliter (range: <400–930,000 copies per milliliter). The most recent on-treatment plasma HIV RNA concentration was <400 copies per milliliter in 192 (72%) of 265 children tested (median duration on treatment: >3 years). Seventy-two children did not have a recent plasma HIV RNA concentration available for inclusion in the data set.

The mean baseline CD4+ lymphocyte count for the 299 children, who have both a baseline count and ≥1 follow-up count obtained ≥1 year after initiation of lopinavir/ritonavir-containing HAART was 292 cells per microliter (range: 1–1433 cells per microliter). The mean change from baseline in CD4+ lymphocyte count was 266 (n = 284), +317 (n = 260), +343 (n = 176), and +270 cells per microliter (n = 121) after 1, 2, 3, and 4 years of treatment, respectively (for all comparisons, P < .0001).

DISCUSSION

We report the long-term follow-up of 414 children and adolescents who received lopinavir/ritonavir-containing HAART at the Romanian-American Children’s Center in Constanta between November 2001 and August 2006. We believe that this may be the largest reported long-term experience with the use of HAART of any kind for HIV-infected children and adolescents living in a low- or middle-income country. In addition, it may represent the world’s largest experience with the use of this particular drug in children.

We treated children and adolescents with HAART by prescription, in a routine pediatric outpatient setting, without the benefit of antiretroviral drug resistance testing to guide therapeutic decisions. Treatment history alone dictated which antiretroviral drugs were given concomitantly with lopinavir/ritonavir. Many children who previously had received zidovudine and lamivudine were treated with stavudine and didanosine, a combination of nucleoside reverse transcriptase inhibitors rarely used today.

Most of the children and adolescents that we treated with lopinavir/ritonavir-containing HAART had been treated previously with other antiretroviral drugs (including HIV protease inhibitors) and had moderate or severe HIV-associated clinical symptoms and immunosuppression. These characteristics often are associated with treatment benefit less pronounced than that typically observed in children naïve to antiretroviral therapy and with less advanced HIV disease. Nevertheless, lopinavir/ritonavir-containing HAART was safe and effective. Eighty-one percent of the children remained on therapy after a median duration of ≥4 years. The mortality rate that we observed in this cohort compares favorably with the rate observed among HIV-infected children living in Constanta between 1998 and 2002. Seventy-two percent of children tested had plasma HIV RNA concentrations <400 copies per milliliter after a median duration of treatment of >3 years, and marked CD4+ lymphocyte count increases were observed after 1, 2, 3, and 4 years of treatment. Comparable virus load and CD4+ lymphocyte count findings were reported from a 48-week multicenter clinical trial.
of 100 children treated with lopinavir/ritonavir-containing HAART.³

Because of the unique epidemiology of pediatric HIV infection in Romania,⁴ the children and adolescents that we treated with lopinavir/ritonavir-containing HAART were relatively old (mean age: 13 years), and most had acquired the infection horizontally. We are limited in our ability to comment on the generalizability of our findings to infants and young children with vertical HIV infection. Nevertheless, lopinavir/ritonavir-containing HAART seems to be a safe, effective, and durable treatment option for antiretroviral drug–experienced older children and adolescents with advanced HIV disease. HAART can be administered safely and effectively to children and adolescents in resource-limited settings.

ACKNOWLEDGMENTS
This work was supported in part by the Abbott Fund’s Step Forward program, the Fogarty International Center of the National Institutes of Health (Baylor AIDS International Training and Research Program grant 5 D43 TW01036), and by the National Institute of Allergy and Infectious Diseases (Baylor Center for AIDS Research grant AI36211).

REFERENCES
Brain Abnormalities in Patients With Hyperimmunoglobulin E Syndrome

Alexandra F. Freeman, MDa, Christina J. Collura-Burke, DO, MPHb, Nicholas J. Patronas, MDc, Lidia Stana Ilcus, MDd, Dirk Darnell, MSNa, Joie Davis, APRN, APNGd, Jennifer M. Puck, MDc, Steven M. Holland, MDa

aNational Institute of Allergy and Infectious Diseases, cClinical Center, and dNational Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland; bWalter Reed Army Medical Center, Washington, DC; eDepartment of Pediatrics, University of California, San Francisco, California

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

ABSTRACT

OBJECTIVES. Hyperimmunoglobulin E syndrome is a multisystem disorder with abnormalities of the immunologic, connective tissue, and skeletal tissue systems. Central nervous system abnormalities have not been considered a feature of hyperimmunoglobulin E syndrome. We aimed to determine whether central nervous system abnormalities detected on brain MRI exist in hyperimmunoglobulin E syndrome and to characterize any identified abnormalities.

PATIENTS AND METHODS. Fifty patients aged from 3 to 52 years (mean: 24 years) with established diagnoses of hyperimmunoglobulin E syndrome had MRI of the brain as part of an hyperimmunoglobulin E syndrome natural history protocol. Abnormalities were described, measured, counted, and mapped. Patient charts were reviewed for neurologic findings and blood pressure measurements.

RESULTS. Focal brain lesions exhibiting high signal intensities on flow-attenuated inversion recovery and on T2-weighted techniques were found in 35 of the 50 patients. The focal hyperintensities were predominantly in the white matter of the cerebral hemispheres, and the number ranged from 2 to >50. The hyperintensities occurred more frequently in adults than in children, and no association with elevated blood pressure was found. Five patients had lacunar infarctions. Chiari type 1 malformations were found in 9 of 50 patients. Two patients had infectious complications presenting on MRI as cerebritis in 1 patient and as a hemorrhagic infarct in the other; both were found on autopsy to be fungal. Neurologic abnormalities were present in 1 patient with a lacunar infarction, the 2 patients with infectious complications, and in 1 patient with a subarachnoid hemorrhage secondary to a berry aneurysm.

CONCLUSIONS. Central nervous system abnormalities are common in hyperimmunoglobulin E syndrome. Focal T2 hyperintensities, not appreciated previously, represent a prominent feature of this rare disease that may assist in diagnosis. The etiology and clinical implications of these abnormalities remain to be investigated.
HYPERIMMUNOGLOBULIN E SYNDROME (HIES; or Job’s syndrome) is a primary immunodeficiency characterized by recurrent skin and lung infections, eczema, and extreme elevations of serum immunoglobulin E. In addition to the immunologic abnormalities, there are connective tissue and morphologic abnormalities in many patients, including characteristic facies, scoliosis, joint hyperextensibility, retained primary dentition, craniosynostosis, osteopenia, and pathologic fractures. Family pedigrees may show an autosomal dominant pattern of inheritance with variable expressivity, but most cases are sporadic. A distinct autosomal recessive form of HIES lacks the skeletal and dental features.

Central nervous system (CNS) abnormalities in patients with HIES have not been described frequently. In our previous review of 30 HIES patients, 3 had noninfectious vascular events in the central nervous system, including an occlusion of the central retinal artery, bilateral aneurysms at the internal carotid bifurcation, an ischemic stroke in the territory of the left middle cerebral artery, and another stroke in the left posterior inferior cerebellar artery. These events prompted a study of cerebral MRI findings in a larger cohort of HIES patients to assess the prevalence, extent, and nature of CNS abnormalities.

METHODS

Fifty patients with HIES were evaluated under an institutional review board-approved natural history protocol at the National Institutes of Health between 1998 and 2004. Study subjects were diagnosed with HIES by experienced clinicians, assisted by a diagnostic scoring system, in which a score of 40 is consistent with HIES, a score of 20 to 40 is indeterminate, and a score of <20 is not suggestive of HIES. Individuals were included in this study if the HIES score was >40 or >20 with a first-degree relative with a score >40.

We obtained brain MRI scans as a part of a prospective study of the natural history of HIES. We have also included in this analysis MRI studies obtained for clinical care before enrollment. Conventional T2-weighted (repetition time [TR]/echo time [TE]: 3500–4850 milliseconds/99–104 milliseconds) and T1-weighted (TR/TE: 400–450 milliseconds/8–15 milliseconds) images were obtained in a 1.5-T magnet. In all but 3 patients, a fluid-attenuated inversion recovery (FLAIR) scan was also obtained using the following imaging parameters: TR/TE: 10 002 milliseconds/145–148 milliseconds, slice thickness of 5 mm, 22 mL field of view, and a matrix of 192 × 256. Postcontrast studies were performed in a subset of patients when needed for clinical management but were not included in this study. Twenty-seven of the 50 patients had >1 MRI scan with an interval of 2 to 6 years (mean: 4.5 years). A total of 133 MRI scans was evaluated.

Each study was reviewed by 1 neuroradiologist (Dr Patronas) and abnormalities recorded. The number of brain lesions found on FLAIR or on the T2-weighted scans was counted, and the size and location of each lesion in the brain parenchyma (white versus gray matter) was recorded. In patients with >1 study, changes in the number of lesions on subsequent scans were noted. Patients with downward displacement of the cerebellar tonsils were classified as Chiari type 1 malformation if the lower pole of the tonsils projected >4 mm below a line connecting the tip of the clivus to the occipital bone.

Charts were reviewed for blood pressure measurements and neurologic deficits. Three blood pressure measurements spanning ≥2 years were averaged for all of the patients ≥18 years, and ≥2 blood pressures spanning ≥2 years were averaged for patients <18 years old.

RESULTS

Fifty patients with HIES, 23 males and 27 females, had brain MRIs. Their ages ranged from 3 to 52 years (mean: 24 years). The HIES scores in this cohort of patients ranged from 29 to 100 (mean: 69). Four of the patients had HIES scores <40; 3 of these patients were <10 years old and, therefore, may have had less assigned points from the scoring system because they had less time to accumulate as many infections, fractures, and retained teeth, and the characteristic facial appearance may not have yet appeared.

Thirty-five (70%) of 50 patients had focal punctate lesions in the white matter, which demonstrated high signal intensity on FLAIR and on T2-weighted sequences (Fig 1 and Table 1). The majority of lesions involved the subcortical and deep white matter, whereas the periventricular white matter was relatively spared. The incidence of such hyperintensities was greater in adults than children, with 9 (47%) of 19 patients ≤18 years and 25 (81%) of 31 patients ≥19 years having focal hyperintensities (Fig 2). The youngest patient with focal hyperintensities was 4 years old. The number of lesions per patient ranged from 2 to 50. Distribution between males 16 (70%) of 23 and females 19 (70%) of 27 was the same. Distribution of HIES scores was 29 to 100 (mean: 71) for patients with hyperintensities and 31 to 87 (mean: 63) for those without.

In the 27 patients who had >1 MRI scan, an increase of ≥3 lesions was observed in 6 (22%) of 27. The mean increase in the number of lesions was 11.5. The age of patients with an increase in the number of lesions ranged from 25 to 49 years (median: 31 years), whereas the age of the patients without a change in lesion number ranged from 3 to 52 years (median: 25 years). In no patient was there a decrease in the number of lesions on follow-up scans, and the location of the lesions remained stable over time.

In 10 (20% overall) of the 35 patients with white matter abnormalities, gray matter lesions were found.
Five patients had punctate lesions, 2 in cortical gray matter and 3 in the basal ganglia. The other 5 patients had focal hyperintense lesions in the basal ganglia presenting as lacunar infarctions (Fig 3). Lacunar infarctions were noted at relatively young ages (32, 33, 37, 38, and 40 years; mean: 36 years). In 3 of these patients, the lacunar infarction was identified incidentally on the initial MRI scan in the absence of neurologic symptoms. In 2 patients, the infarct was new compared with previous scans. One was a 37-year-old man with acute onset of ataxia and ophthalmoplegia, whereas the other, a 33-year-old woman, was neurologically intact.

For the 25 patients >18 years old, patients with brain lesions did not have more hypertension than those without brain lesions. The mean blood pressure for patients with brain lesions was 121/79 mm Hg and for those without it was 132/79 mm Hg. Thirty-two percent of patients (8 of 25) with focal brain lesions had diastolic blood pressures >80 mm Hg, and 28% (3 of 25) had systolic pressures >130 mm Hg, as compared with 33% of patients (2 of 6) without such lesions with diastolic pressures >80 mm Hg and 67% of patients (4 of 6) with systolic pressures >130 mm Hg.

Two patients had infectious complications identified on MRI and confirmed at autopsy. Fatal *Scedosporium prolificans* cerebritis developed in the setting of extensive *Scedosporium* bronchitis and pneumonia in a 24-year-old woman. A fatal mycotic aneurysm because of *Aspergillus fumigatus* developed in the setting of a modest pulmonary aspergiloma in a 29-year-old woman. The mycotic aneurysm developed in the left middle cerebral artery circulation and presented on MRI as a hemorrhagic infarct and clinically as acute loss of consciousness with seizures.

Nine (18%) of the 50 patients had type 1 Chiari malformations (Fig 4). Bilateral internal carotid bifurcation berry aneurysms with subarachnoid hemorrhage were found in 1 patient; a venous angioma of the frontal lobe was found in another patient; capillary telangiectasia of the pons was found in 1 patient; arachnoid cysts were found in 2 patients; and third ventricle colloid cyst was found in 1 patient. With the exception of the subarachnoid hemorrhage complicating the berry aneurysms, all of these findings were incidental abnormalities in individuals without neurologic abnormalities.

**DISCUSSION**

HIES is a multisystem disorder with abnormalities of the immunologic, connective tissue, and skeletal systems. There are few previous reports of CNS abnormalities in sporadic or autosomal dominant HIES, mostly related to infections. Recently, an autosomal recessive variant of HIES has been described that lacks many of the multisystem features of typical HIES but has a high incidence of severe viral infections, CNS vasculitis, and autoimmune abnormalities. The great majority of abnormalities that we found in autosomal dominant or sporadic HIES were without neurologic abnormalities, ranging from focal hyperintensities and Chiari 1 malformations to lacunar infarcts and CNS infections.

Focal hyperintense brain lesions were found in 70% of the patients, a rate much higher than the 0.5% rate in a study of 1000 healthy volunteers (mean age: 30.6 years; range: 3–83 years). Focal hyperintensities are also found as an incidental finding in the elderly and in neurofibromatosis type 1 (NF1). Focal brain hyper-
intense lesions are very common in elderly individuals, with a prevalence reaching >90% in some studies.\textsuperscript{11,15} These lesions are indistinguishable from the abnormalities encountered in our patients. In the elderly, the hyperintense lesions have been associated with increased blood pressures, evidence of previous silent stroke, smoking, and other vascular risk factors.\textsuperscript{15,17} They are known to represent manifestations of small vessel disease in various stages of evolution from perivascular edema secondary to ischemia at 1 end of the spectrum to lacunar or cortical ischemic infarction in the other. Published reports indicate that these lesions may be associated with cognitive decline.\textsuperscript{17,18} The white matter hyperintensities and cortical and deep gray matter lesions in our patients with HIES were found at ages much younger than would be expected in the general population, suggesting that there may be small vessel disease not recognized previously. Although we did not find significant differences in blood pressure between the HIES patients with and without focal brain lesions, the numbers are limited because of the rarity of HIES. Continuous ambulatory blood pressure monitoring may be indicated to observe a significant difference. Whether cognitive decline occurs in HIES, especially when there is an increase in lesion load over time, needs additional evaluation.

The hyperintensities described in NF1 seem very dif-
ferent from those in HIES and are likely of a different etiology. Hyperintensities in NF1 are typically located in deep gray structures (such as the basal ganglia), cerebellum, and the brainstem and occur more frequently in children than adults.\textsuperscript{12,13,16} The etiology of these lesions in NF1 remains unknown, although it has been hypothesized that they result from intramyelinic edema.\textsuperscript{19}

Chiari 1 malformation was observed in 18\% of HIES patients. There is limited information of the incidence of asymptomatic Chiari 1 malformation in the general population. However, in a retrospective study examining 22 591 brain MRIs, Chiari 1 malformation (defined as $>$5 mm of tonsillar herniation below the foramen magnum) was found in only 165 studies (0.77\%).\textsuperscript{20} Although their criteria were more stringent than ours in defining the degree of herniation (5 mm of tonsillar herniation versus 4 mm), the percentage of patients that we observed with tonsillar herniation significantly exceeded theirs. Because Chiari 1 malformation may be associated with syringomyelia, careful neurologic histories and examinations are necessary.

The finding of disseminated fungal infection to the brain in 2 patients emphasizes that fungal pulmonary infections, generally superinfections of pneumatocele cavities, in these patients need to be aggressively treated. Screening with brain MRI may be indicated in patients with pulmonary fungal disease even in the absence of any neurologic changes. Patients with HIES are typically associated with syringomyelia. The Rotterdam Scan Study.\textsuperscript{19}

CNS abnormalities on brain MRI are a remarkably common and previously unrecognized aspect of HIES. Several patients with lacunar infarcts in addition to focal hyperintensities suggest possible small vessel disease. Additional prospective neurologic and neuropsychologic evaluation is necessary to determine the clinical significance, if any, of these lesions in HIES. These newly recognized CNS abnormalities confirm the complex multisystem nature of HIES. These lesions are frequent enough in HIES and uncommon enough in the general population, especially at early ages, to be considered in future diagnostic classifications, and brain MRI may be useful in the diagnosis of uncertain cases.

CONCLUSIONS

CNS abnormalities on brain MRI are a remarkably common and previously unrecognized aspect of HIES. Several patients with lacunar infarcts in addition to focal hyperintensities suggest possible small vessel disease. Additional prospective neurologic and neuropsychologic evaluation is necessary to determine the clinical significance, if any, of these lesions in HIES. These newly recognized CNS abnormalities confirm the complex multisystem nature of HIES. These lesions are frequent enough in HIES and uncommon enough in the general population, especially at early ages, to be considered in future diagnostic classifications, and brain MRI may be useful in the diagnosis of uncertain cases.

ACKNOWLEDGMENTS

This research was supported by the Intramural Research Programs of the National Institutes of Allergy and Infectious Diseases, the Clinical Center, and the National Human Genome Research Institute at the National Institutes of Health.

REFERENCES

10. Katzman GL, Dagher AP, Patronas NJ. Incidental findings on brain magnetic resonance imaging from 1000 asymptomatic volunteers. JAMA. 1999;282:36–39
A Longitudinal Study of the Prevalence, Development, and Persistence of HIV/Sexually Transmitted Infection Risk Behaviors in Delinquent Youth: Implications for Health Care in the Community

Erin Gregory Romero, BS*, Linda A. Teplin, PhD*, Gary M. McClelland, PhD*, Karen M. Abram, PhD*, Leah J. Welty, PhD*, Jason J. Washburn, PhD, ABPP*

Departments of *Psychiatry and Behavioral Sciences and **Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois

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

ABSTRACT

OBJECTIVES. Our goal was to examine the prevalence, development, and persistence of drug and sex risk behaviors that place delinquent youth at risk for HIV and other sexually transmitted infections.

METHODS. At the baseline interview, HIV/sexually transmitted infection drug and sex risk behaviors were assessed in a stratified random sample of 800 juvenile detainees aged 10 to 18 years. Participants were reinterviewed approximately 3 years later. The final sample in these analyses (n = 724) included 316 females and 408 males; there were 393 African American participants, 198 Hispanic participants, 131 non-Hispanic white participants, and 2 participants who self-identified their race as “other.”

RESULTS. More than 60% of youth had engaged in ≥10 risk behaviors at their baseline interview, and nearly two thirds of them persisted in ≥10 risk behaviors at follow-up. Among youth living in the community, many behaviors were more prevalent at follow-up than at baseline. Among incarcerated youth, the opposite pattern prevailed. Compared with females, males had higher prevalence rates of many HIV/sexually transmitted infection risk behaviors and were more likely to persist in some behaviors and develop new ones. Yet, injection risk behaviors were more prevalent among females than males and were also more likely to develop and persist. Overall, there were few racial and ethnic differences in patterns of HIV/sexually transmitted infection risk behaviors; most involved the initiation and persistence of substance use among non-Hispanic whites and Hispanics.

CONCLUSIONS. Because detained youth have a median stay of only 2 weeks, HIV/sexually transmitted infection risk behaviors in delinquent youth are a community public health problem, not just a problem for the juvenile justice system. Improv-
ing the coordination among systems that provide HIV/sexually transmitted infection interventions to youth—primary care, education, mental health, and juvenile justice—can reduce the prevalence of risk behaviors and substantially reduce the spread of HIV/sexually transmitted infection in young people.

![Image](https://via.placeholder.com/150)

**ADOLESCENTS AND YOUNG adults are disproportionately affected by HIV and other sexually transmitted infections (STIs).** Youth aged 15 to 24 years represent approximately 25% of sexually active persons in the United States but accounted for nearly 50% of new STI cases (9.1 million) in 2000. Between 2001 and 2005, HIV/AIDS diagnoses increased >20% in persons aged 13 to 24 years. Advances for treating AIDS have slowed mortality. Still, among persons aged 25 to 34 years, HIV is the sixth leading cause of death among non-Hispanic whites and Hispanics, the third leading cause of death among African Americans, and the leading cause of death in African American women.

HIV/AIDS and other STIs are increasingly diseases of racial/ethnic minorities and youth. The National Longitudinal Study of Adolescent Health, which sampled >13,000 young adults, found that the rate of HIV infection in African Americans was 4.9 cases per 1000 persons, compared with 0.22 cases per 1000 in other racial/ethnic groups. The most recent statistics compiled by the Centers for Disease Control and Prevention indicate that more than three quarters of persons <25 years diagnosed with HIV/AIDS are African American or Hispanic. Young minority females are at particular risk. African American and Hispanic females account for approximately 80% of HIV/AIDS diagnoses in females aged 13 to 24 years. Minorities have greater exposure to risk factors than do other groups, including low socioeconomic status, urban living, substance abuse, and limited access to health care.

Minorities are also overrepresented in the juvenile justice system, where HIV/STI risk behaviors are prevalent. Detained youth report more risk behaviors and initiate them at younger ages than do youth in the community. Detained youth are likely to be at continued risk for HIV infection as they age. Adults in prison have higher rates of HIV/STI risk behaviors and HIV infection (1.8%) than the general population (0.2%). Sound public policy and effective interventions require data on the developmental course of HIV/STI risk behaviors. Because youth are detained for an average of only 2 weeks, their behaviors place persons in the community at risk.

There are, however, few comprehensive studies of HIV/STI risk behaviors in delinquent youth, and, to our knowledge, no longitudinal studies. Even after expanding our literature review to include “high-risk” youth, such as inner city youth and other impoverished populations, we found only 4 epidemiological studies with follow-up periods >6 months. Only 1 of these studies collected comprehensive information on HIV/STI sex and drug risk behaviors. None of these studies investigated how the development and persistence of HIV/STI risk behaviors differ by gender, race/ethnicity, and age.

To our knowledge, this is the first large-scale longitudinal study of HIV/STI risk behaviors in delinquent youth. Our study has 2 methodologic strengths: a stratified random sample, large enough \((n = 724)\) to generate reliable rates of HIV/STI risk behaviors for key demographic subgroups (eg, females and Hispanics) and comprehensive measures of HIV/STI drug and sex risk behaviors.

In this article, we address 3 questions:

- Prevalence: among youth in the sample, what proportion reported each HIV/STI behavior?
- Development: among youth who did not report a specific HIV/STI behavior at baseline, what proportion reported that behavior at follow-up?
- Persistence: among youth who did report a specific HIV/STI behavior at baseline, what proportion persisted in that behavior at follow-up?

We examine differences according to incarceration status and demographic variables (gender, race/ethnicity, and age).

**METHODS**

**Sampling Procedures**

Our data are from the Northwestern Juvenile Project, a longitudinal study of health needs and outcomes of delinquent youth. We recruited a stratified random sample of 1829 detained youth initially arrested and detained awaiting the adjudication and/or disposition of their case between November 20, 1995, and June 14, 1998, at the Cook County Juvenile Temporary Detention Center (CCJTD) in Chicago, IL. To ensure adequate representation of key subgroups, we stratified our sample by age (10–13 years or ≥14 years), gender, race/ethnicity (African American, non-Hispanic white, and Hispanic), and legal status (processed as a juvenile or an adult). The CCJTD is used for pretrial detention and for offenders sentenced for <30 days. Consistent with juvenile detainees nationwide, >80% of detainees at CCJTD were male, and most were racial/ethnic minorities. Additional information on our methods has been published elsewhere.

**Procedures to Obtain Assent and Consent**

This research was approved by the institutional review boards of Northwestern University, the Centers for Disease Control and Prevention, and the US Office of Pro-
tection from Research Risks. At the baseline and follow-up interviews, participants signed either an assent form (if they were <18 years) or a consent form (if they were ≥18 years). The Northwestern University Institutional Review Board and the Centers for Disease Control and Prevention Institutional Review Board waived parental consent, consistent with federal regulations regarding research with minimal risk (45 Code of Federal Regulations [CFR] 46.116[c], 45 CFR 46.116[d], and 45 CFR 46.408[c]). We nevertheless tried to contact parents to provide them information and offer an opportunity to decline participation. Despite repeated attempts to contact the parent or guardian, for 43.8% of the participants, none could be found. In lieu of parental permission to contact the parent or guardian, for 43.8% of the participants, none could be found. In lieu of parental consent, an independent participant advocate representing the interests of the participants oversaw youth assent. Federal regulations allow for a participant advocate if parental consent is not feasible (45 CFR 46.116[d]).

Participants
Collection of the baseline HIV/STI data began when funding became available, from February 1997 through June 1998. Among the 1052 youth sampled during this period, 3.9% (n = 41) refused to participate. There were no significant differences in refusal rates according to gender, race/ethnicity, or age. Fourteen participants did not complete the HIV/STI questions because of the interviewer’s error. One participant was released from detention before finishing the interview; 196 participants left the detention center while we were locating their caretakers to obtain consent or before we could schedule an interview. The final number of youth who received the HIV/STI interview was 800; of these, 769 (96.1%) were interviewed at follow-up; 12 (1.5%) died before the follow-up; 3 (0.4%) withdrew from the study; and 16 (2.0%) were lost to follow-up. Time to follow-up was between 2.9 and 7.9 years (mean [SD] follow-up: 3.3 [0.6] years; median follow-up: 3.1 years).

Fifty-five of the 769 participants were excluded from our analyses: 5 (0.7%) did not receive the HIV/STI risk behavior assessment at follow-up (because of time constraints or the interviewer’s error); and 40 (5.2%) received their follow-up interview >4.5 years after their baseline interview. We chose 4.5 years for the cutoff, because, in this high-risk and highly mobile sample, participants can be difficult to track; using a stricter cutoff would restrict the generalizability of the sample. To ensure that our cutoff did not bias the findings, we compared the demographic characteristics (gender, race/ethnicity, and age) of participants who were interviewed between 3.5 and 4.5 years (n = 81; 11% of the sample) after baseline with those interviewed within 3.5 years after baseline; there were no significant differences. In addition, we examined whether our findings were affected by including these participants. We repeated all of the analyses using only participants interviewed within 3.5 years; the findings were substantially the same.

The final sample in these analyses (n = 724) included 316 females and 408 males; there were 393 African American participants, 198 Hispanic participants, 131 non-Hispanic white participants, and 2 participants who self-identified their race as “other.” At baseline, 113 youth were processed as adults, and 611 were processed as juveniles. The median length of stay at CCJRTC was 15 days (range: 1–686 days; mean [SD] days: 40.7 [75.3]). At baseline, participants were aged 10 to 18 years (mean [SD] age: 14.8 [1.4]; median age: 15). At follow-up, participants were aged 13 to 22 years (mean [SD] age: 18.1 [1.4]; median age: 18). Time to follow-up was 2.9 to 4.5 years (mean [SD] time to follow-up: 3.2 [0.3] years; median: 3.1 years).

Procedures for Data Collection
At the baseline interview, face-to-face, structured interviews were conducted at the detention center in a private area; most interviews took place within 2 days of intake. At the follow-up, the same participants were interviewed, irrespective of where they lived. Participants were interviewed in the community (66.2%), at correctional facilities (26.2%), at residential placement facilities (2.5%), or by telephone if they lived in a community >2 hours away (5.1%). Baseline and follow-up interviews took 2 to 4 hours to complete. We used both male and female interviewers; female interviewers were interviewed only by female interviewers. Most interviewers had advanced degrees in psychology or an associated field and had experience interviewing at-risk youth. All of the interviewers were trained for ≥1 month by 1 of the authors (Dr Abram) and other supervisory staff. One third of the interviewers were fluent in Spanish.

Measures
We examined behaviors associated with increased risk for HIV/STI, including sex risk behaviors and injection risk behaviors (sharing needles or “works” for drug injection, piercings, or tattoos). We also examined antecedents to HIV/STI risk behaviors, such as alcohol and other drug use, because they may indirectly lead to HIV/STIs by increasing high-risk sexual behaviors.

HIV/STI risk behaviors were assessed using the National Institute on Drug Abuse (NIDA) Risk Behavior Assessment (RBA). Although designed for adults, we chose the RBA because instruments designed for adolescents and young adults did not assess the breadth, frequency, and severity of HIV/AIDS risk behaviors common in our sample. A report issued by the Substance Abuse and Mental Health Services Administration recommends the RBA for the comprehensive assessment of HIV/AIDS risk behaviors among drug-using adolescents. The RBA is a reliable and valid measure of drug
We supplemented the RBA with items from the Adolescent Health Survey from NIDA’s Study of Street Youth at Risk for AIDS and Yale’s AIDS Risk Inventory. Experts reviewed our measure at baseline, and we pilot-tested 58 participants. At baseline, lifetime drug use was assessed using screen items (1 for each substance) from the Diagnostic Interview Schedule 2.3. At baseline and follow-up, recency and frequency of drug use were assessed using NIDA’s RBA.

Missing Data

Missing Cases
To assess the effect of attrition on generalizability, we compared participants who provided follow-up data with those who did not on the following variables: demographic characteristics (gender, race/ethnicity, and age) and HIV/STI risk behaviors reported at baseline. There were no significant differences except those who died were more likely to be male ($P < .05$), and those lost to follow-up were more likely to be non-Hispanic white or Hispanic ($P < .05$) and were less likely to have had sex with >1 partner ($P < .05$). Potential bias from demographic differences in attrition was adjusted by weighting the statistical analyses by sampling strata (see “Statistical Analysis” in “Methods”).

Missing Data From Interviews Conducted by Telephone
Because telephone interviews needed to be shorter than face-to-face interviews, they are missing the following variables at follow-up ($n = 37; 5.1\%$): use of specific drugs, types of sex with a high-risk partner, sex and unprotected sex while drunk or high, and trading sex and drugs. Comparing participants interviewed by telephone with those interviewed face-to-face revealed the following: (1) no significant demographic differences (gender, race/ethnicity, or age); (2) no significant differences in the prevalence of HIV/STI risk behaviors reported at baseline; and (3) no significant differences in the prevalence of other HIV/STI risk behaviors reported at follow-up.

Independent Variables
We compared HIV/STI risk behaviors by gender, race/ethnicity, and age. We also examined incarceration status since baseline. For behaviors assessed “since the last interview,” participants were considered incarcerated if they self-reported that they had been “mostly in correctional facilities” since the baseline interview (21.4\% of sample; 126 males and 29 females). For behaviors assessed “in the past 3 months” or less, participants were considered incarcerated if they self-reported that they had been “mostly in correctional facilities in the past 3 months” (23.1\% of sample; 138 males and 29 females).

Statistical Analysis
All of the data were weighted to reflect the population at the CCJTDC. Because selected strata were oversampled, we used sample weights, based on CCJTDC’s population, to estimate descriptive statistics and model parameters that reflect CCJTDC’s population. Taylor series linearization was used to estimate SEs. Only statistically significant findings with $P < .05$ are noted in the text.

Changes in the prevalence of behaviors between the baseline and follow-up interviews were assessed using paired differences with an adjusted Wald F statistic. Logistic regression was used to assess demographic differences in the prevalence (Tables 1 and 2), development (Tables 3 and 4), and persistence (Tables 5 and 6) of individual risk behaviors. The independent variables in the regression models were incarceration status only (Tables 1 and 2), incarceration status and gender (Tables 2, 3, and 5), incarceration status and race/ethnicity (Tables 4 and 6), and incarceration status and age. We tested for differences between specific groups (eg, African American versus Hispanic) only when the overall model was significant at the $P < .05$ level. We controlled for incarceration status in all of the analyses by either computing separate prevalence rates for those incarcerated and those in the community or including incarceration status in logistic regression models.

RESULTS

Prevalence of HIV/STI Risk Behaviors

Comparing the Baseline and Follow-up Interviews

Males
Table 1 shows that prevalence of the following behaviors increased at follow-up: oral sex, anal sex (receptive and/or insertive), sex while drunk or high, and unprotected sex while drunk or high. In contrast, multiple sex partners (>1 and >3 in the past 3 months), recent use of marijuana, and frequent use of marijuana decreased at follow-up.

Table 1 also shows differences according to incarceration status. Among males in the community, most behaviors were more prevalent at follow-up. Only 2 behaviors, recent and frequent use of marijuana, were significantly less prevalent. Among incarcerated males, the opposite pattern prevailed; many behaviors were less prevalent at follow-up. Only 1 behavior, oral sex with a high-risk partner, was significantly more prevalent.

Females
Table 2 shows that the prevalence of the following behaviors increased at follow-up: sexual activity, vaginal sex, recent unprotected vaginal sex, oral sex, recent unprotected oral sex, unprotected sex while drunk or high, and trading sex and drugs. In contrast, multiple sex partners (>1 in the past 3 months), use of alcohol, use of
<table>
<thead>
<tr>
<th>Variable</th>
<th>Total Males (n = 408)</th>
<th>Living Situation at Follow-up*</th>
<th>Significant Difference Between Incarcerated and Community at Follow-up, P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline Prevalence, %</td>
<td>Follow-up Prevalence, %</td>
<td>Significance, P</td>
</tr>
<tr>
<td>Sex risk behaviors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexually active</td>
<td>90.6 92.9</td>
<td>90.1 99.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Multiple partners: &gt;1 in past 3 mo</td>
<td>61.7 63.8</td>
<td>&lt;.001</td>
<td>63.7 53.4</td>
</tr>
<tr>
<td>Multiple partners: &gt;3 in past 3 mo</td>
<td>35.2 15.5</td>
<td>&lt;.01</td>
<td>33.5 22.2</td>
</tr>
<tr>
<td>Vaginal sex</td>
<td>90.3 92.3</td>
<td>NS</td>
<td>89.6 98.6</td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected vaginal sex</td>
<td>36.4 44.4</td>
<td>NS</td>
<td>35.1 62.2</td>
</tr>
<tr>
<td>Vaginal sex with high-risk partnerb</td>
<td>25.1 37.9</td>
<td>NS</td>
<td>21.7 36.6</td>
</tr>
<tr>
<td>Oral sex</td>
<td>43.1 58.8</td>
<td>&lt;.05</td>
<td>39.5 63.2</td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected oral sex</td>
<td>32.9 35.1</td>
<td>NS</td>
<td>32.1 47.9</td>
</tr>
<tr>
<td>Oral sex with high-risk partnerb</td>
<td>8.1 16.3</td>
<td>NS</td>
<td>11.4 13.1</td>
</tr>
<tr>
<td>Anal sex (receptive and/or insertive)</td>
<td>9.7 22.3</td>
<td>&lt;.05</td>
<td>9.8 17.3</td>
</tr>
<tr>
<td>Receptive anal sex</td>
<td>1.6 0.5</td>
<td>NS</td>
<td>2.3 0.7</td>
</tr>
<tr>
<td>Insertive anal sex (males only)</td>
<td>9.7 22.1</td>
<td>NS</td>
<td>9.8 17.0</td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected anal sex</td>
<td>3.7 7.1</td>
<td>NS</td>
<td>4.1 10.5</td>
</tr>
<tr>
<td>Anal sex with high-risk partnerb</td>
<td>25.6 63.3</td>
<td>NS</td>
<td>35.1 16</td>
</tr>
<tr>
<td>Sex while drunk or high</td>
<td>65.8 80.2</td>
<td>&lt;.05</td>
<td>64.6 85.0</td>
</tr>
<tr>
<td>Unprotected sex while drunk or high</td>
<td>36.0 54.9</td>
<td>&lt;.01</td>
<td>33.6 53.8</td>
</tr>
<tr>
<td>Traded sex and drugs</td>
<td>28.6 68.3</td>
<td>NS</td>
<td>1.2 9.9</td>
</tr>
<tr>
<td>Drug and injection risk behaviors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used alcohol</td>
<td>89.2 89.1</td>
<td>NS</td>
<td>88.0 91.5</td>
</tr>
<tr>
<td>Recent (past month) use of alcohol</td>
<td>56.7 47.8</td>
<td>NS</td>
<td>62.0 64.7</td>
</tr>
<tr>
<td>Frequent use of alcohol (&gt;3 times past month)</td>
<td>30.8 29.1</td>
<td>NS</td>
<td>31.5 38.4</td>
</tr>
<tr>
<td>Used marijuana</td>
<td>93.4 88.5</td>
<td>NS</td>
<td>93.8 96.1</td>
</tr>
<tr>
<td>Recent (past month) use of marijuana</td>
<td>77.5 46.1</td>
<td>&lt;.001</td>
<td>84.9 62.7</td>
</tr>
<tr>
<td>Frequent use of marijuana (&gt;3 times past month)</td>
<td>60.4 37.4</td>
<td>&lt;.001</td>
<td>67.0 50.0</td>
</tr>
<tr>
<td>Used other substancec</td>
<td>14.2 15.1</td>
<td>NS</td>
<td>17.2 17.9</td>
</tr>
<tr>
<td>Recent (past month) use of other substancec</td>
<td>7.0 3.7</td>
<td>NS</td>
<td>8.8 4.7</td>
</tr>
<tr>
<td>Frequent use of other substance (&gt;3 times past month)c</td>
<td>1.4 1.2</td>
<td>NS</td>
<td>1.7 1.2</td>
</tr>
<tr>
<td>Injected drugs</td>
<td>0.1 0.1</td>
<td>NS</td>
<td>0.2 0.1</td>
</tr>
<tr>
<td>Tattooed</td>
<td>43.8 43.8</td>
<td>NS</td>
<td>39.5 41.8</td>
</tr>
<tr>
<td>Shared needle(s) or equipment (injection drug use/tattooing)</td>
<td>3.5 0.3</td>
<td>NS</td>
<td>4.9 0.4</td>
</tr>
<tr>
<td>Shared needle(s) in a risky location (injection drug use/tattooing)</td>
<td>1.6 0.0</td>
<td>NS</td>
<td>2.3 0.0</td>
</tr>
<tr>
<td>Shared needle(s) without cleaning (injection drug use/tattooing)</td>
<td>1.6 0.0</td>
<td>NS</td>
<td>2.3 0.0</td>
</tr>
</tbody>
</table>

NS indicates not significant. This table uses all of the cases that have baseline and follow-up data for each variable. Follow-up data are measured “since the last interview” unless noted. Data are weighted to reflect the actual population of the CCJTDC.

a For variables measured “since the last interview,” the number of “incarcerated participants” is composed of those who self-reported that they had “been mostly in correctional facilities” since the baseline interview (126 males and 29 females). For variables measured “in the last 3 months” or less, the number of incarcerated participants is composed of those that self-reported they had been “mostly in correctional facilities in the past 3 months” (138 males and 29 females).
b “High-risk partners” include persons who have ever worked as a prostitute, persons with HIV/AIDS, persons who inject drugs, and persons whose sexual history is not well known.
c “Other substance” includes substances other than alcohol or marijuana.
d Tests of significance could not be computed, because baseline and follow-up prevalence rates were 0.
e Tests of significance could not be computed because ≥1 cell size was 0.
f “Risky locations” include a park, street, alley, abandoned building, car, public bathroom, and crack house/shooting gallery.
Table 2: Prevalence of HIV/STI Sex and Drug Risk Behaviors in Female Juvenile Detainees at the Baseline and Follow-up Interviews (n = 316)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total Females (n = 316)</th>
<th>Living Situation at Follow-up*</th>
<th>Significant Difference Between Incarcerated and Community at Follow-up, P</th>
<th>Odds Ratios (95% Confidence Intervals) for Follow-up Prevalence, Males to Females (Adjusted for Incarceration Status)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Follow-up</td>
<td>Significant Change, P</td>
<td>Baseline</td>
</tr>
<tr>
<td>Sex-risk behaviors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexually active</td>
<td>870</td>
<td>946</td>
<td>&lt;.001</td>
<td>864</td>
</tr>
<tr>
<td>Multiple partners: &gt;1 in past 3 mo</td>
<td>269</td>
<td>133</td>
<td>&lt;.001</td>
<td>262</td>
</tr>
<tr>
<td>Multiple partners: &gt;3 in past 3 mo</td>
<td>51</td>
<td>23</td>
<td>NS</td>
<td>48</td>
</tr>
<tr>
<td>Vaginal sex</td>
<td>837</td>
<td>940</td>
<td>&lt;.001</td>
<td>831</td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected vaginal sex</td>
<td>508</td>
<td>61.9</td>
<td>&lt;.01</td>
<td>509</td>
</tr>
<tr>
<td>Vaginal sex with high-risk partner&lt;sup&gt;d&lt;/sup&gt;</td>
<td>192</td>
<td>25.4</td>
<td>NS</td>
<td>195</td>
</tr>
<tr>
<td>Oral sex</td>
<td>316</td>
<td>48.4</td>
<td>&lt;.001</td>
<td>331</td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected oral sex</td>
<td>237</td>
<td>35.9</td>
<td>&lt;.01</td>
<td>243</td>
</tr>
<tr>
<td>Oral sex with high-risk partner&lt;sup&gt;d&lt;/sup&gt;</td>
<td>34</td>
<td>6.3</td>
<td>NS</td>
<td>37</td>
</tr>
<tr>
<td>Anal sex (rectal and/or insertive)</td>
<td>77</td>
<td>9.8</td>
<td>NS</td>
<td>84</td>
</tr>
<tr>
<td>Anal sex (rectal)</td>
<td>77</td>
<td>9.8</td>
<td>NS</td>
<td>84</td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected anal sex</td>
<td>18</td>
<td>3.6</td>
<td>NS</td>
<td>19</td>
</tr>
<tr>
<td>Anal sex with high-risk partner&lt;sup&gt;d&lt;/sup&gt;</td>
<td>35</td>
<td>0.3</td>
<td>NS</td>
<td>38</td>
</tr>
<tr>
<td>Sex while drunk or high</td>
<td>530</td>
<td>58.7</td>
<td>NS</td>
<td>522</td>
</tr>
<tr>
<td>Unprotected sex while drunk or high</td>
<td>346</td>
<td>45.7</td>
<td>&lt;.01</td>
<td>341</td>
</tr>
<tr>
<td>Sexed sex and drugs</td>
<td>31</td>
<td>8.0</td>
<td>&lt;.01</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Drug and injection risk behaviors:

| Used alcohol | 909 | 85.6 | <.05 | 908 | 86.8 | NS | 92.2 | 743 | NS | NS | 1.7 (0.7–4.1) |
| Recent (past month) use of alcohol | 537 | 48.8 | NS | 520 | 52.2 | NS | 72.1 | 140 | <.01 | <.001 | 1.6 (0.9–3.0) |
| Frequent use of alcohol (>3 times past month) | 263 | 23.1 | NS | 246 | 24.9 | NS | 44.5 | 38 | <.01 | <.05 | 1.9 (1.1–3.4) |
| Used marijuana | 904 | 81.9 | <.001 | 89.8 | 82.2 | <.01 | 96.3 | 798 | NS | NS | 3.7 (2.2–6.2) |
| Recent (past month) use of marijuana | 674 | 52.3 | <.001 | 660 | 56.8 | <.01 | 82.2 | 63 | <.01 | <.001 | 1.3 (0.7–2.3) |
| Frequent use of marijuana (>3 times past month) | 442 | 38.0 | NS | 43.3 | 41.4 | NS | 53.5 | 38 | <.01 | <.01 | 1.5 (0.8–2.6) |
| Used other substance<sup>g</sup> | 215 | 24.6 | NS | 204 | 23.2 | NS | 32.3 | 388 | NS | NS | 0.6 (0.4–1.1) |
| Recent (past month) use of other substance<sup>g</sup> | 84 | 5.1 | NS | 7.8 | 5.4 | NS | 14.8 | 25 | NS | NS | 0.9 (0.5–1.6) |
| Frequent use of other substance (>3 times past month)<sup>g</sup> | 27 | 1.9 | NS | 1.7 | 2.1 | NS | 12.3 | 0 | <.05 | <.01 | 0.6 (0.2–1.9) |
| Injected drugs | 10 | 0.9 | NS | 0.7 | 1.0 | NS | 3.8 | 0 | NS | NS | 0.1 (0.01–0.9) |
| Tattooed | 462 | 42.0 | NS | 464 | 41.9 | NS | 54.0 | 435 | NS | NS | 1.0 (0.6–1.7) |
| Shared needle(s) or equipment (injection drug use/tattooing) | 13 | 1.0 | NS | 1.1 | 0.8 | NS | 3.0 | 30 | NS | NS | 0.3 (0.05–2.3) |
| Shared needle(s) in a risky location (injection drug use/tattooing)<sup>h</sup> | 0.0 | 0 | NS | 0.0 | 0 | NS | 0.0 | 0 | NS | NS | — |
| Shared needle(s) without cleaning (injection drug use/tattooing)<sup>h</sup> | 0.0 | 0 | NS | 0.0 | 0 | NS | 0.0 | 0 | NS | NS | — |

NS indicates not significant. This table uses all of the cases that have baseline and follow-up data for each variable. Follow-up data are measured “since the last interview” unless noted. Data are weighted to reflect the actual population of the CCJTC.

<sup>a</sup>For variables measured “since the last interview,” the number of incarcerated participants is composed of those who self-reported they had “been mostly in correctional facilities” since the baseline interview (126 males and 29 females). For variables measured “in the last 3 months” or less, the number of incarcerated participants is composed of those who self-reported they had been “mostly in correctional facilities in the past 3 months” (138 males and 29 females).

<sup>b</sup>P < .05.

<sup>c</sup>P < .01.

<sup>d</sup>“High-risk partners” include persons who have ever worked as a prostitute, persons with HIV/AIDS, persons who inject drugs, and persons whose sexual history is not well known.

<sup>e</sup>Tests of significance could not be computed because baseline and follow-up prevalence rates were 0.

<sup>f</sup>Tests of significance could not be computed because ≥1 cell size was 0.

<sup>g</sup>“Other substance” includes substances other than alcohol or marijuana.

<sup>h</sup>“Risky locations” include a park, street, alley, abandoned building, car, public bathroom, and crack house/shooting gallery.
marijuana, and recent use of marijuana were less prevalent at follow-up.

Table 2 also reports differences according to incarceration status. Among females in the community, many behaviors were more prevalent at follow-up. Only 4 behaviors, multiple sex partners (>1 in past 3 mo), use of marijuana, and recent use of marijuana, were significantly less prevalent. Among incarcerated females, the opposite pattern prevailed; only 1 behavior, oral sex, was significantly more prevalent.

Prevalence at Follow-up

Males

Table 1 shows that at follow-up, nearly all of the males were sexually active, had vaginal sex, and used alcohol and marijuana. The prevalence of unprotected
### TABLE 4: Development of HIV/STI Sex and Drug Risk Behaviors in Juvenile Detainees Between the Baseline and Follow-up Interviews According to Race/Ethnicity (n = 722)

<table>
<thead>
<tr>
<th>Variable</th>
<th>African American, %</th>
<th>Non-Hispanic White, %</th>
<th>Hispanic, %</th>
<th>Overall Test of Racial/Ethnic Differences for Developing (Adjusted for Incarceration Status), ( P )</th>
<th>Odds Ratios (95% Confidence Intervals) for Developing (Adjusted for Incarceration Status)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((n = 393))</td>
<td>((n = 131))</td>
<td>((n = 198))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex risk behaviors</td>
<td></td>
<td></td>
<td></td>
<td>( \chi^2 ) ( \text{(df = 1)} )</td>
<td></td>
</tr>
<tr>
<td>Sexually active</td>
<td>7.9 93.1</td>
<td>21.1 78.9</td>
<td>14.0 86.0</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Multiple partners: &gt;1 in past 3 mo</td>
<td>37.0 92.9</td>
<td>64.3 90.0</td>
<td>49.9 94.9</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Multiple partners: &gt;3 in past 3 mo</td>
<td>82.5 12.0</td>
<td>85.2 3.8</td>
<td>79.8 6.9</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Vaginal sex</td>
<td>8.5 88.3</td>
<td>21.1 78.9</td>
<td>14.0 94.9</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected vaginal sex</td>
<td>64.1 41.8</td>
<td>62.6 44.8</td>
<td>58.1 283</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Vaginal sex with high-risk partner</td>
<td>74.5 33.6</td>
<td>84.6 43.0</td>
<td>76.6 389</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Oral sex</td>
<td>57.4 42.6</td>
<td>47.2 64.5</td>
<td>64.1 39.8</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected oral sex</td>
<td>66.8 23.1</td>
<td>60.1 45.2</td>
<td>75.9 218</td>
<td>&lt;0.05</td>
<td>0.4 (0.2–1.0) ( e ) 0.9 (0.3–2.3) 2.1 (1.0–4.5)</td>
</tr>
<tr>
<td>Oral sex with high-risk partner</td>
<td>90.7 12.5</td>
<td>93.6 24.6</td>
<td>98.4 169</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Anal sex (rectal and/or insertive)</td>
<td>91.0 19.2</td>
<td>89.6 16.9</td>
<td>89.5 334</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Anorectal anal sex</td>
<td>97.8 0.8</td>
<td>98.7 2.9</td>
<td>99.9 19</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Insertive anal sex (males only)</td>
<td>90.7 20.0</td>
<td>89.6 16.5</td>
<td>88.8 355</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected anal sex</td>
<td>97.0 5.0</td>
<td>94.1 4.4</td>
<td>97.5 161</td>
<td>&lt;0.05</td>
<td>1.4 (0.3–7.0) 0.2 (0.04–1.2) 0.2 (0.03–0.7) ( e )</td>
</tr>
<tr>
<td>Anal sex with high-risk partner</td>
<td>97.4 5.8</td>
<td>97.0 4.0</td>
<td>98.9 72</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Sex while drunk or high</td>
<td>33.3 66.7</td>
<td>38.0 60.6</td>
<td>42.6 57.5</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Unprotected sex while drunk or high</td>
<td>66.3 43.7</td>
<td>62.0 42.7</td>
<td>58.0 41.1</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Traded sex and drugs</td>
<td>96.9 3.1</td>
<td>98.5 4.5</td>
<td>98.0 3.3</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Drug and injection risk behaviors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used alcohol</td>
<td>13.2 86.8</td>
<td>5.2 60.0</td>
<td>2.3 669</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Recent (past month) use of alcohol</td>
<td>44.9 48.8</td>
<td>37.4 41.5</td>
<td>290 640</td>
<td>&lt;0.001</td>
<td>0.2 (0.1–0.5) ( e ) 0.1 (0.03–0.5) ( e ) 0.6 (0.2–1.2)</td>
</tr>
<tr>
<td>Frequent use of alcohol (&gt;3 times past month)</td>
<td>69.8 21.2</td>
<td>73.4 38.9</td>
<td>67.7 416</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Used marijuana</td>
<td>5.9 70.6</td>
<td>6.5 27.4</td>
<td>10.8 110</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Recent (past month) use of marijuana</td>
<td>23.4 24.8</td>
<td>22.9 38.0</td>
<td>23.5 234</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Frequent use of marijuana (&gt;3 times past month)</td>
<td>39.1 25.7</td>
<td>38.3 38.3</td>
<td>48.9 270</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Used other substance</td>
<td>97.2 2.8</td>
<td>41.4 22.2</td>
<td>50.4 247</td>
<td>&lt;0.001</td>
<td>0.02 (0.01–0.04) ( e ) 0.04 (0.02–0.08) ( e ) 2.5 (1.0–6.4)</td>
</tr>
<tr>
<td>Recent (past month) use of other substance</td>
<td>97.8 2.0</td>
<td>75.2 41.5</td>
<td>79.9 28</td>
<td>&lt;0.001</td>
<td>0.02 (0.01–0.07) ( e ) 0.04 (0.01–0.13) ( e ) 16 (0.6–4.4)</td>
</tr>
<tr>
<td>Frequent use of other substance (&gt;3 times past month)</td>
<td>99.9 0.1</td>
<td>71.5 43.3</td>
<td>74.5 24</td>
<td>&lt;0.001</td>
<td>0.01 (0.01–0.29) ( e ) 0.02 (0.02–0.7) ( e ) 20 (0.4–10.2)</td>
</tr>
<tr>
<td>Injected drugs</td>
<td>100.0 0.0</td>
<td>97.6 2.3</td>
<td>99.8 0.0</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Tattooed</td>
<td>99.9 27.4</td>
<td>61.7 35.8</td>
<td>37.9 316</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Shared needle(s) or equipment (injection drug use/tattooing)</td>
<td>96.0 4.0</td>
<td>99.7 1.7</td>
<td>98.4 11</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Shared needle(s) in a risky location (injection drug use/tattooing)</td>
<td>98.1 1.9</td>
<td>100.0 0.0</td>
<td>100.0 0.0</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Shared needle(s) without cleaning (injection drug use/tattooing)</td>
<td>98.1 1.0</td>
<td>1000.0 0.0</td>
<td>999.9 0.1</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

NS indicates not significant. One male and 1 female who self-identified as "other" race/ethnicity were not included in this analysis. This table uses all of the cases that have baseline and follow-up data for each variable. Follow-up data are measured "since the last interview" unless noted. Data are weighted to reflect the actual population of the CCJTDC.

\( a \) "Development" is defined as follows: among youth who did not report a specific HIV/STI behavior at baseline, what proportion reported that behavior at follow-up?

\( b \) We tested for differences between specific groups (eg, African American versus Hispanic) only when the overall model was significant at the <0.05 level.

\( c \) Tests of significance are adjusted for incarceration status (yes/no). For variables measured "since the last interview," participants are considered incarcerated if they self-reported they had "been mostly in correctional facilities" since the baseline interview (126 males and 29 females). For variables measured "in the last 3 months" or less, participants are considered incarcerated if they self-reported that they had been "mostly in correctional facilities in the past 3 months" (138 males and 29 females).

\( d \) "High-risk partners" include persons who have ever worked as a prostitute, persons with HIV/AIDS, persons who inject drugs, and persons whose sexual history is not well known.

\( e \) Odds ratios and tests of significance could not be computed because ≥1 cell size was 0.

\( f \) "Other substances" includes substances other than alcohol or marijuana.

\( g \) "Other substance" includes substances other than alcohol or marijuana.

\( h \) Odds ratios and tests of significance could not be computed because ≥1 cell size was 0.

\( i \) "Risky locations" include a park, street, alley, abandoned building, car, public bathroom, and crack house/shooting gallery.
sex in the past 3 months was also notable: nearly half had recent unprotected vaginal sex, more than one third had recent unprotected oral sex, and 7% had recent unprotected anal sex, and more than half had unprotected sex while drunk or high.

Table 1 also shows that, at follow-up, more males in the community than incarcerated males reported the following behaviors: sexually active, multiple sex partners (>1 and >3 in the past 3 months), vaginal sex, recent unprotected vaginal sex, recent unprotected oral sex, recent unprotected anal sex, traded sex and drugs, recent use of alcohol, frequent use of alcohol, use of marijuana, frequent use of marijuana, and frequent use of marijuana. In contrast, more incarcerated males engaged in anal sex with a high-risk partner than those in the community.

### Females

Table 2 shows that, at follow-up, nearly all of the females were sexually active and had vaginal sex; more than four fifths used alcohol and marijuana. The preva-
## TABLE 6: Persistence of HIV/STI Sex and Drug Risk Behaviors in Juvenile Detainees Between the Baseline and Follow-up Interviews According to Race/Ethnicity (n = 722)

<table>
<thead>
<tr>
<th>Variable</th>
<th>African American, % (n=393)</th>
<th>Non-Hispanic White, % (n=131)</th>
<th>Hispanic, % (n=198)</th>
<th>Overall Test of Racial/Ethnic Differences for Persisting (Adjusted for Incarceration Status), P&lt;.05</th>
<th>Odds Ratios (95% Confidence Intervals) for Developing (Adjusted for Incarceration Status)*, P&lt;.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Absent at Baseline</td>
<td>% Developing</td>
<td>% Absent at Baseline</td>
<td>% Developing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex risk behaviors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexually active</td>
<td>92.1</td>
<td>94.2</td>
<td>78.9</td>
<td>96.0</td>
<td>86.0</td>
</tr>
<tr>
<td>Multiple partners: &gt;1 in past 3 mo</td>
<td>63.0</td>
<td>41.9</td>
<td>35.7</td>
<td>34.2</td>
<td>50.1</td>
</tr>
<tr>
<td>Vaginal sex</td>
<td>37.5</td>
<td>24.6</td>
<td>14.8</td>
<td>14.5</td>
<td>20.2</td>
</tr>
<tr>
<td>Unprotected vaginal sex</td>
<td>91.5</td>
<td>93.8</td>
<td>78.9</td>
<td>93.9</td>
<td>85.8</td>
</tr>
<tr>
<td>Vaginal sex with high-risk partner</td>
<td>25.5</td>
<td>40.8</td>
<td>15.4</td>
<td>72.6</td>
<td>23.4</td>
</tr>
<tr>
<td>Oral sex</td>
<td>42.6</td>
<td>60.4</td>
<td>52.8</td>
<td>89.0</td>
<td>35.9</td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected oral sex</td>
<td>33.2</td>
<td>51.8</td>
<td>39.9</td>
<td>84.6</td>
<td>241</td>
</tr>
<tr>
<td>Oral sex with high-risk partner</td>
<td>9.3</td>
<td>35.1</td>
<td>6.4</td>
<td>57.9</td>
<td>1.6</td>
</tr>
<tr>
<td>Sex risk behavior (non-drug)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral sex (non-drug)</td>
<td>9.0</td>
<td>13.6</td>
<td>104</td>
<td>36.7</td>
<td>10.5</td>
</tr>
<tr>
<td>Anus (non-drug)</td>
<td>2.2</td>
<td>19.1</td>
<td>13</td>
<td>50.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Insertion of anus (male only)</td>
<td>9.3</td>
<td>13.5</td>
<td>104</td>
<td>34.9</td>
<td>11.2</td>
</tr>
<tr>
<td>Recent (past 3 mo) unprotected anal sex</td>
<td>3.0</td>
<td>5.0</td>
<td>59</td>
<td>17.9</td>
<td>2.5</td>
</tr>
<tr>
<td>Anal sex with high-risk partner</td>
<td>2.6</td>
<td>4.2</td>
<td>3.0</td>
<td>11.0</td>
<td>1.1</td>
</tr>
<tr>
<td>Sex while drunk or high</td>
<td>65.0</td>
<td>85.0</td>
<td>62.0</td>
<td>89.7</td>
<td>57.4</td>
</tr>
<tr>
<td>Unprotected sex while drunk or high</td>
<td>33.7</td>
<td>70.9</td>
<td>38.0</td>
<td>68.2</td>
<td>42.0</td>
</tr>
<tr>
<td>Drug use and drug use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used alcohol</td>
<td>86.8</td>
<td>88.9</td>
<td>94.8</td>
<td>95.8</td>
<td>97.7</td>
</tr>
<tr>
<td>Recent (past month) use of alcohol</td>
<td>55.1</td>
<td>55.2</td>
<td>62.6</td>
<td>74.5</td>
<td>61.0</td>
</tr>
<tr>
<td>Frequent use of alcohol (&gt;3 times past month)</td>
<td>30.2</td>
<td>29.6</td>
<td>26.6</td>
<td>52.5</td>
<td>32.3</td>
</tr>
<tr>
<td>Used marijuana</td>
<td>94.1</td>
<td>89.8</td>
<td>93.5</td>
<td>92.1</td>
<td>89.2</td>
</tr>
<tr>
<td>Recent (past month) use of marijuana</td>
<td>76.6</td>
<td>49.9</td>
<td>77.1</td>
<td>69.0</td>
<td>76.5</td>
</tr>
<tr>
<td>Frequent use of marijuana (&gt;3 times past month)</td>
<td>60.9</td>
<td>42.4</td>
<td>61.7</td>
<td>61.7</td>
<td>51.1</td>
</tr>
<tr>
<td>Used other substance</td>
<td>2.8</td>
<td>11.4</td>
<td>58.6</td>
<td>86.4</td>
<td>496</td>
</tr>
<tr>
<td>Recent (past month) use of other substance</td>
<td>2.2</td>
<td>5.5</td>
<td>32.1</td>
<td>44.7</td>
<td>20.1</td>
</tr>
<tr>
<td>Frequent use of other substance (&gt;3 times past month)</td>
<td>0.1</td>
<td>0.0</td>
<td>12.5</td>
<td>16.9</td>
<td>4.3</td>
</tr>
<tr>
<td>Injected drugs</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Tattooed</td>
<td>40.1</td>
<td>56.4</td>
<td>38.3</td>
<td>63.4</td>
<td>62.1</td>
</tr>
<tr>
<td>Shared needle(s) or equipment (injection drug use/tattooing)</td>
<td>4.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.6</td>
</tr>
<tr>
<td>Shared needle(s) in a risky location (injection drug use/tattooing)</td>
<td>1.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Shared needle(s) without cleaning (injection drug use/tattooing)</td>
<td>1.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

NS indicates not significant. One male and 1 female who self-identified as “other” race/ethnicity were not included in this analysis. This table uses all of the cases that have baseline and follow-up data for each variable. Follow-up data are measured “since the last interview” unless noted. Data are weighted to reflect the actual population of the CCJTDC.

*Persistence* is defined as follows: among youth who engaged in a behavior at the baseline interview, what proportion persisted in that behavior at the follow-up interview?

b We tested for differences between specific groups (eg, African American versus Hispanic) only when the overall model was significant at the <.05 level.

c Tests of significant are adjusted for incarceration status (yes/no). For variables measured “since the last interview,” participants are considered incarcerated if they self-reported they had “been mostly in correctional facilities” since the baseline interview (126 males and 29 females). For variables measured “in the last 3 months” or less, participants are considered incarcerated if they self-reported they had been “mostly in correctional facilities in the past 3 months” (138 males and 29 females).

d “High-risk partners” include persons who have ever worked as a prostitute, persons with HIV/AIDS, persons who inject drugs, and persons whose sexual history is not well known.

*P* < .05.

f Odds ratios and tests of significance could not be computed because 1 cell size was 0.

g “Other substance” includes substances other than alcohol or marijuana.

h *P* < .001.

i *P* < .01.

j “Risky locations” include a park, street, alley, abandoned building, car, public bathroom, and crack house/shooting gallery.
lence of unprotected sex in the past 3 months was also notable: nearly two thirds had recent unprotected vaginal sex, more than one third had recent unprotected oral sex, nearly 4% had recent unprotected anal sex, and nearly half had unprotected sex while drunk or high.

Table 2 also shows that, at follow-up, more females in the community than incarcerated females reported the following behaviors: sexually active, vaginal sex, recent unprotected vaginal sex, recent use of alcohol, frequent use of alcohol, recent use of marijuana, and frequent use of marijuana.

Gender Differences

There were many gender differences, which are reported in Table 2. More males than females reported the following behaviors: multiple sex partners (>1 and ≥3 in the past 3 months), vaginal sex, oral sex with a high-risk partner, anal sex (receptive and/or insertive), anal sex with a high-risk partner, sex while drunk or high, frequent use of alcohol, and use of marijuana. In contrast, more females than males reported the following behaviors: receptive anal sex and injection drugs.

Age Differences (Data Not Shown)

More youth ≥18 years (n = 502) than youth <18 years (n = 222) reported the following behaviors: recent unprotected anal sex (prevalence rate: 2.0% vs 7.9%; P < .05), recent use of alcohol (prevalence rate: 28.9% vs 52.9%; P < .05), and tattooing (prevalence rate: 19.8% vs 50.1%; P < .001).

Development of HIV/STI Risk Behaviors

Table 3 shows that many participants who had not reported risk behaviors at baseline had developed them by follow-up. For example, among those who had not previously reported unprotected vaginal sex, nearly 40% of males and more than half of females reported such behavior at follow-up. More than 40% of males and nearly 40% of females began engaging in unprotected sex while drunk or high at follow-up.

Gender Differences

Table 3 shows that, at follow-up, more males than females had begun engaging in the following behaviors: multiple sex partners (>1 and ≥3 in the past 3 months), vaginal sex with a high-risk partner, anal sex (receptive and/or insertive), anal sex with a high-risk partner, use of alcohol, and frequent use of alcohol. In contrast, at follow-up more females than males had begun engaging in receptive anal sex, use of substances other than alcohol and marijuana, and injection drugs.

Racial/Ethnic Differences

Table 4 shows that there were few racial and ethnic differences in the development of risk behaviors; most involved the initiation of substance use among non-Hispanic whites and Hispanics. More non-Hispanic whites and Hispanics than African Americans had begun engaging in recent use of alcohol, use of substances other than alcohol or marijuana, recent use of substances other than alcohol or marijuana, and frequent use of substances other than alcohol or marijuana. More non-Hispanic whites than African Americans had begun having recent unprotected oral sex. More Hispanics than non-Hispanic whites had begun having recent unprotected anal sex.

Persistence of HIV/STI Risk Behaviors

Table 5 shows that persistence of sex and drug risk behaviors was common for both males and females. For example, among youth who had engaged in unprotected vaginal sex at baseline, more than half of males and nearly 70% of females persisted in this behavior at follow-up. More than three quarters of males and nearly 60% of females persisted in unprotected sex while drunk or high at follow-up. More than 70% of males and nearly 70% of females persisted in using substances other than alcohol and marijuana at follow-up.

Gender Differences

There were few gender differences in the tendency for risk behaviors to persist. At follow-up, more males than females had persisted in the following sex risk behaviors: multiple sex partners (>1 in the past 3 months), sex while drunk or high, and use of marijuana.

Racial/Ethnic Differences

Table 6 shows that there were few racial or ethnic differences in the tendency for behaviors to persist. More non-Hispanic whites and Hispanics than African Americans persisted in the following behaviors: oral sex and use of substances other than alcohol and marijuana. More Hispanics than African Americans persisted in recent use of alcohol and tattooing.

DISCUSSION

Our findings show that youth involved in the juvenile justice system continue to be at great risk for HIV/STIs as they age. Nearly three quarters of youth engaged in ≥1 unprotected sexual risk behavior at follow-up. More than 60% had engaged in ≥10 risk behaviors at their baseline interview, and nearly two thirds of them persisted in ≥10 risk behaviors at follow-up.

Irrespective of gender, race/ethnicity, or age, sex risk behaviors were prevalent and likely to persist and develop. At follow-up, more than one third of males and one fourth of females reported engaging in vaginal sex with a high-risk partner. At baseline, more than one third of males and more than half of females reported engaging in recent unprotected vaginal sex. At follow-up, more than half of youth persisted in this behavior, and more than one third developed this behavior. These findings underscore the importance of providing early
HIV/STI interventions, continued outreach, and long-term interventions that focus on sex risk behaviors.

Infection risk behaviors were uncommon at baseline and at follow-up. However, our findings on risk behaviors related to noninjection drug use are of great concern. One half of our participants had a substance use disorder at baseline, and >80% of youth reported using alcohol and marijuana at follow-up. At baseline, more than one third of participants engaged in unprotected sex while drunk or high. At follow-up, approximately three fifths of the youth persisted in this behavior, and two fifths of the youth developed this behavior. Substance abuse can lead to high-risk sexual behaviors by affecting decision-making, compromising judgment, decreasing the likelihood of condom use, and increasing the likelihood of sex-for-drug exchanges and injection drug use. Yet, research on noninjection drug use and HIV/STIs has lagged, considering its importance in the current HIV/STI epidemic.

Taken together, these findings mirror the changing patterns of transmission of HIV/STIs in the general population. In the early stages of the HIV/STI epidemic, the most common patterns of transmission were injection drug use (approximately one quarter of AIDS cases) and male-to-female sex (two thirds of AIDS cases). Male-to-female sexual contact at that time accounted for only 4% of AIDS cases; it now accounts for one third of reported HIV/AIDS cases. We found a number of gender differences, even after adjusting for incarceration status. Compared with females, males had higher prevalence rates of many HIV/STI risk behaviors and were more likely to persist in some behaviors and develop new ones. Yet, injection risk behaviors were more prevalent among females than males and were also more likely to develop and persist. Our findings emphasize the need to develop interventions tailored to specific patterns of risk and transmission. For example, nearly 80% of females contract HIV/AIDS from vaginal sex compared with only 16% of males. Gender-specific interventions are especially important now that females compose nearly 20% of juvenile detainees and 26% of AIDS cases (compared with 7% during the early years of the epidemic).

Incarceration status was an important variable. Among youth in the community, many behaviors were more prevalent at follow-up than at baseline. Among incarcerated youth, the opposite pattern prevailed. Our findings add to the growing debate on the role of incarceration in the HIV/STI epidemic. One view is that correctional facilities are “breeding grounds” for HIV/AIDS. Others suggest that the disproportionately high prevalence of HIV/AIDS in correctional facilities occurs because behaviors that put persons at risk for HIV/AIDS (eg. drug use, prostitution) also put them at risk for incarceration. Although risk behaviors may be less common in correctional facilities than in the community, they may carry substantially greater risk. For example, to prevent HIV/STI transmission, prisoners may use plastic gloves and hand lotion instead of lubricated condoms. Similarly, to inject drugs, inmates may share needles or “works” or use dirty equipment if sterilization is unavailable. Moreover, the probability of infection is also higher, because more persons in prisons than in the community are infected with HIV.

Overall, there were few racial and ethnic differences in patterns of HIV/STI risk behaviors; most involved the initiation and persistence of substance use among non-Hispanic whites and Hispanics. There were surprisingly few racial/ethnic differences in sex risk behaviors. Yet, because of the disproportionate numbers of African Americans who cycle through correctional facilities, the pediatrics community must focus on implementing culturally appropriate interventions for African American youth and young adults. More than any other racial/ethnic group, African Americans are disproportionately incarcerated and affected by HIV/AIDS. Although African Americans compose only 13% of the general population, and juvenile crime rates are relatively similar across race/ethnicity, African Americans compose about 40% of incarcerated youth and adults and 49% of new cases of HIV/AIDS (52% among men and 66% among women).

The prevalence of HIV/STI risk behaviors in our sample is similar to that of other high-risk youth: those living on the street, drug users, and those living in the inner city. Primary risk reduction interventions may not reach these youth. Although most schools now provide HIV/STI education, youth who are frequently truant, such as delinquent and homeless youth, are unlikely to receive school-based interventions. Moreover, delinquent youth are overrepresented in groups that are uninsured (including the poor, youth living in central cities, and older adolescents), reducing the likelihood that they will have a primary care physician from whom they could receive primary interventions. Public clinics and emergency services are often the primary source of health care for high-risk youth. As recently recommended by the American Academy of Pediatrics, public clinics should integrate HIV prevention, especially sex education and substance abuse treatment, into primary medical care. HIV/STI interventions should also be provided in detention centers and in juvenile courts that, based on recent statistics, could reach as many as 1.6 million youth annually.

Limitations

It was not possible to assess actual HIV/STI risk behaviors, such as having unprotected sex with an infected partner and sharing injection/piercing equipment with an infected partner. Moreover, it was not feasible to obtain biological outcome measures, such as HIV or STI
tests. Thus, our measures of HIV/STI risk are proximal. Findings might have been slightly different had follow-up data been available for participants who died, withdrew from the study, or were lost to follow-up. We examined HIV/STI risk behaviors during 2 periods of our subjects’ lives. Our analyses do not address causal mechanisms underlying HIV/STI risk. Our findings, drawn from 1 site, may pertain only to youth who were detained during adolescence in urban detention centers of similar demographic composition. Our sample (though larger than most previous investigations) limited our analyses of demographic subgroups that are less common in detention centers, such as young, non-Hispanic white females. Finally, the data are subject to the limitations of self-reporting. Participants may have underreported some behaviors and exaggerated others.

CONCLUSIONS
HIV infection and disproportionate minority confinement are among the most critical racial/ethnic disparities in our nation. Future studies must explore the causes and correlates of the increased risk of HIV/STIs faced by minority youth. To develop appropriate public policy initiatives, future studies must also disentangle the role of incarceration in the transmission of HIV/STIs. Because most detained youth return to their communities, HIV/STI risk behaviors in delinquent youth are a community public health problem, not just a problem for the juvenile justice system. Improving the coordination among systems that provide HIV/STI interventions to youth—primary care, education, mental health, and juvenile justice—can reduce the prevalence of risk behaviors and substantially reduce the spread of HIV/STIs in young people. The Surgeon General will soon issue a “Call to Action on Correctional Health.” By targeting HIV/STI risk behaviors in delinquent youth, we have the opportunity to redress significant health disparities and threats to public health.

ACKNOWLEDGMENTS
This work was supported by National Institute of Mental Health grants R01MH54197 and R01MH59463 (Division of Services and Intervention Research and Center for Mental Health Research on AIDS) and Office of Juvenile Justice and Delinquency Prevention grants 1999-JE-FX-1001 and 2005-JL-FX-0288. Major funding was also provided by the National Institute on Drug Abuse, the Substance Abuse and Mental Health Services Administration (Center for Mental Health Services, Center for Substance Abuse Prevention, and Center for Substance Abuse Treatment), the National Institutes of Health Center on Minority Health and Health Disparities, the Centers for Disease Control and Prevention (National Center on Injury Prevention and Control and National Center for HIV, STD, and TB Prevention), the National Institute on Alcohol Abuse and Alcoholism, the National Institutes of Health Office of Research on Women’s Health, the National Institutes of Health Office on Rare Diseases, Department of Labor, the William T. Grant Foundation, and the Robert Wood Johnson Foundation. Additional funds were provided by the John D. and Catherine T. MacArthur Foundation, the Open Society Institute, and the Chicago Community Trust. We thank all of our agencies for their collaborative spirit and steadfast support.

Many more people than the authors contributed to this project. This study could not have been accomplished without the advice of Ann Hohmann, PhD, Kimberly Hoagwood, PhD, Heather Ringeisen, PhD, and Eugene Griffin, PhD. Jacques Normand, PhD, Helen Cesari, MS, Richard Needle, PhD, Robert Booth, PhD, David Huizinga, PhD, and David Ostrow, MD, PhD, generously offered their expertise in developing our instruments. David Stoff, PhD, Grayson Norquist, MD, and Delores Parron, PhD, provided support and encouragement. Cellia Fisher, PhD, guided our human subject procedures. We thank Frank Palella, Jr, MD, and the anonymous reviewers for their many helpful comments. We also thank all of the project staff, especially Amy Mericle, PhD, for instrumentation development and project leadership and Lynda Carey, MA, for her superb data management. We thank Kate Elkington, PhD, for literature review. We also greatly appreciate the cooperation of everyone working in the Cook County and State of Illinois systems. Without their cooperation, this study would not have been possible. Finally, we thank the participants for their time and willingness to participate and our field interviewers for their commitment to excellence.

REFERENCES


85. Stiffman AR, Dore P, Cunningham RM. Inner-city youths and


93. Dembo R. Problems among youths entering the juvenile justice system, their service needs and innovative approaches to address them. *Subst Use Misuse.* 1996;31:81–94


97. Short PF, Graefe DR. Battery-powered health insurance? stability in coverage of the uninsured. *Health Aff (Millwood).* 2003;22:244–255


ARTICLE

Hospital-Based Directly Observed Therapy for HIV-Infected Children and Adolescents to Assess Adherence to Antiretroviral Medications

Daniel Glikman, MD, Linda Walsh, NP, Judy Valkenburg, PA-C, P. Daisy Mangat, RN, MPH, John F. Marcinak, MD

Pediatric and Adolescent HIV Program, Section of Infectious Diseases, Department of Pediatrics, University of Chicago, Chicago, Illinois

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

ABSTRACT

BACKGROUND. The introduction of highly active antiretroviral therapy for HIV led to significant declines in HIV-associated morbidity and mortality in children. Nonadherence to antiretroviral therapy is the leading cause of treatment failure in HIV-infected patients. The ability to recognize nonadherence is suboptimal, and differentiating it from other causes of inadequate viral suppression may be difficult.

OBJECTIVES. The purpose of this work was to examine the efficacy of hospital-based directly observed therapy in assessing adherence to antiretroviral medications in HIV-infected children and adolescents suspected of nonadherence and failing other interventions.

METHODS. The medical charts of all HIV-infected patients admitted to the University of Chicago Comer Children’s Hospital for directly observed therapy from July 2004 to June 2006 were reviewed. Patients were hospitalized for 7 days. Data collected included demographics, clinical and immune class category, previous and current antiretroviral medications, viral resistance tests, HIV-1 RNA viral load, and CD4+ T-cell number and percentage before and after directly observed therapy.

RESULTS. There were 9 perinatally infected patients with a total of 13 admissions. The median age was 13 years, and 8 had been treated with multiple antiretroviral regimens. Three common patterns of changes in the viral load over time were observed. In the first, the viral load dropped at the end of the directly observed therapy period and stayed low thereafter. In the second, the drop in the viral load seen at the end of the period was not sustained. In the third, there was no change in the viral load during or after the directly observed therapy period. Compared with the viral load at admission, the viral load at the end of directly observed therapy was lower in 8 patients with a mean ± SD decrease of 0.8 ± 0.55 log10 copies per mL.

CONCLUSIONS. Short, hospital-based directly observed therapy was helpful in confirming nonadherence to antiretroviral medications, therefore impacting future therapeutic decisions in HIV-infected children and adolescents. Short, hospital-based directly observed therapy should be considered in patients with poor virological control for whom outpatient interventions have failed.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2614
doi:10.1542/peds.2006-2614
Ms Mangat’s current affiliation is Department of Pediatrics, Section of Infectious Diseases, University of Illinois at Chicago, Chicago, IL.

Key Words
HIV, adherence, directly observed therapy, children, adolescents

Abbreviations
HAART—highly active antiretroviral therapy
DOT—directly observed therapy
VL—viral load

Accepted for publication Nov 2, 2006
Address correspondence to Daniel Glikman, MD, Section of Pediatric Infectious Diseases, Wyler Pavilion, 5841 S Maryland Ave, MC 6054, Chicago, IL 60637. E-mail: dglikman@peds.bsd.uchicago.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
Various reports have documented that highly active antiretroviral therapy (HAART) can inhibit HIV replication and result in significant declines in HIV-associated morbidity and mortality in children and adults. Inadequate suppression of viral replication by HAART can result from poor adherence to therapy, low potency of the antiretroviral regimen, viral resistance to antiretroviral medications, and pharmacokinetic interactions causing inadequate drug delivery.

Nonadherence is the leading cause of HAART failure in suppressing viral replication. It is essential to achieve ≥95% adherence to the HAART regimen to suppress viral replication and avoid the emergence of resistance. HAART, therefore, requires an especially high level of adherence for an indefinite time period to achieve optimal viral suppression.

The problem of nonadherence is significant in the pediatric population because of a lack of liquid formulations for some drugs, the often-required large volume of medications, the extremely poor palatability of some medications, and the dependence of the child on a caregiver to administer the drugs. The caregiver is often also HIV infected, thus further complicating the psychosocial conditions. Nondisclosure of the diagnosis of HIV to the child or family members can further complicate therapy and increase nonadherence. Adolescent patients present even more challenges, given the unique developmental, psychosocial, and lifestyle issues implicit in adolescence.

In a child or adolescent not responding to HAART, nonadherence should be a major consideration. However, the ability to recognize nonadherence is suboptimal, and differentiating nonadherence from other causes of inadequate viral suppression may be difficult, especially in heavily drug-experienced patients with a multidrug resistant virus. Interventions to encourage adherence include patient education and counseling, behavioral contracts, pill boxes, electronic pill monitors, simplified drug regimens, gastrostomy tube placement, reminder calls, home health nurse visits, and directly observed therapy (DOT).

DOT is a direct method to measure adherence and is considered to be the most accurate. The principle of DOT has its roots in the treatment of tuberculosis, for which DOT programs have dramatically improved cure rates. Studies with DOT in HIV-infected patients included mainly community-based interventions for marginalized populations, such as adults with drug abuse and HIV-infected inmates being released from prisons. Hospital-based DOT for HIV-infected children and adolescents has been rarely reported. These studies, however, provide evidence for the potential to substantially decrease the HIV viral load (VL) in a short hospital stay.

We report our experience with hospital-based DOT for HIV-infected children and adolescents cared for by the Pediatric and Adolescent HIV Program at the University of Chicago Comer Children’s Hospital. Since 1990, HIV-infected and exposed infants, as well as HIV-infected children and adolescents, have been cared for by the program; some have been followed for >10 years. We currently care for 73 HIV-infected patients ranging in age from 1 to 23 years and 37 HIV-exposed infants. We used a short hospital-based DOT in 9 patients as an intensive strategy to ensure adherence, differentiate nonadherence from viral drug resistance, look for and address social problems, and evaluate for toxicities of treatment. The objective of this retrospective analysis is to show that hospital-based DOT is an effective approach for managing suboptimal response to HAART in the HIV-infected pediatric and adolescent population.

METHODS

Subjects
The medical charts of all of the HIV-infected children and adolescents admitted to the University of Chicago Comer Children’s Hospital for DOT from July 2004 (the starting date for our inpatient-based DOT policy) until June 2006 were reviewed retrospectively. Data collected from each patient’s record included date of birth, ethnicity, gender, diagnosis date of HIV infection, mode of infection, Centers for Disease Control and Prevention clinical and immune class category, number of previous HAART regimens, current HAART medications, previous interventions to enhance adherence to the medication regimen, viral resistance test results, dates of hospitalization, HIV-1 RNA VL and CD4+ T-cell number and percentage before and after DOT, and changes in HAART regimen made after DOT. Data about significant events during DOT, such as drug-related adverse events, missed or late medication administration, and psychosocial issues were also collected. HIV-1 RNA VL was measured by reverse transcription-polymerase chain reaction with a detection range of 400 to 750 000 copies per mL with the standard procedure or 50 to 100 000 copies per mL with the ultrasensitive procedure (Roche Amplicor 1.5; Roche, Branchburg, NJ). All of the data were entered into a Microsoft Access program (Microsoft, Redmond, WA). The study was reviewed and approved by the University of Chicago Institutional Review Board.

Patients hospitalized for DOT were suspected of nonadherence with the treatment regimen. In most cases, each of the following criteria were met: (1) evidence of treatment failure as indicated by a more than threefold (≥0.5 log_{10}) increase in HIV-1 RNA levels in ≥2 consecutive laboratory evaluations; (2) caregiver or patient’s insistence that adherence to HAART is complete; (3) viral genotype and/or phenotype tests indicating viral susceptibility to the patient’s antiretroviral regimen (partial susceptibility in heavily drug-experienced patients); and (4) no decrease in HIV-1 RNA level after intensified
Differential clinic visit schedule and intensive education about the importance of adherence, including health education group meetings, telephone call reminders, and home visits.

The virological parameter of treatment failure was chosen based on the “Guidelines for the Use of Antiretroviral Agents in Pediatric HIV Infection” and on the observation that HIV-1 VL can fluctuate in the range of 0.5 to 0.7 log_{10} copies per mL in patients receiving stable therapy or no treatment for HIV. We hypothesized that patients nonadherent to antiretroviral therapy will demonstrate significant decrease in the HIV-1 VL after a week of hospital-based DOT.

**RESULTS**

**Description of Study Population and Clinical Data**

A total of 13 admissions of 9 patients occurred during the study period. Three patients were admitted more than once (2 or 3 times). In 1 patient, the hospitalization was shortened because of a substantial decrease in VL value drawn on day 1 of the DOT compared with previous values. All of the patients were perinatally infected and were black. There were 7 girls and 2 boys, and the median age was 13 years (range: 7–17 years). In 6 of the 9 patients, the primary caretaker was an HIV-infected mother. Three patients had gastrostomy tubes in place. Patient characteristics are shown in Table 1. All of the patients had received antiretroviral therapy since the diagnosis of HIV. At the time of DOT, 6 patients were being treated with HAART and 3 with mega-HAART. Except for 1, all of the patients were treated with multiple antiretroviral therapy regimens in the past. Five of 7 patients, in whom HIV resistance testing was performed, experienced triple-class resistance mutations.

Therapeutic data are shown in Table 2.

**Virological and Immunologic Data**

VL assays performed at the beginning and end of DOT provided an objective measure of the effect of adherence. Eight patients in 11 DOT admissions had lower VL values at the completion of DOT compared with values on day 1 of DOT, thereby strongly suggesting nonadherence (Table 2). The mean length of time between the pre-DOT visits to the beginning of DOT was 32.2 days (SD: 15.4 days). The mean length of the DOT hospital stay was 6.8 days (SD: 1.3 days), and the mean length of time between day 1 of DOT and the first follow-up appointment was 47.5 days (SD: 20.7 days).

Three common patterns of changes in the log_{10} VL over time were seen and are depicted in Fig 1. In the first (Fig 1A), the patient experienced a rapid decline in VL observed at the end of the DOT period. In this case, nonadherence was confirmed and relayed to the patient and caregiver together with emphasis on the good results achieved by the patient when the medications were actually taken. The decrease in VL was maintained after

**Data Analysis**

For statistical analysis, 5 periods of time around the DOT hospitalization were chosen. These include the pre-DOT clinic visit to day 1 of the DOT period, day 1 of DOT to last day of the DOT period, day 1 of the DOT to the first clinic visit after the DOT period, day 1 of the DOT to the 6-month post-DOT follow-up period, and the pre-DOT clinic visit to the 6-month post-DOT follow-up period. These periods were chosen to examine immediate and sustained effects of DOT. Paired $t$ tests were used to compare changes in VL, CD4$^+$ T-cell count, and CD4$^+$ T-cell percentage over each period of time. The log_{10} transformation of the VL data was taken before analysis. Nonparametric Wilcoxon signed-rank tests were also used, but because the conclusions were similar, only the results from the paired $t$ tests are reported. A $P < .05$ was considered statistically significant. For analysis purposes, VL listed as less than the lower limit of detection was assigned a value 1 less than that amount, and VL greater than the upper limit of detection was assigned a value 1 greater than that amount (ie, HIV-1 RNA values reported as <50 copies per mL were changed to 49 copies per mL, and values >750 000 copies per mL were reported as 750 001 copies per mL). In patients with multiple DOT visits, only the first visit was used in analysis. In addition, if a subject did not have a visit at 6 months after DOT, the nearest follow-up visit after 6 months was used for calculations. All of the statistical analyses were performed using Stata 9 (Stata Corp, College Station, TX).
TABLE 1  Characteristics of HIV-Infected Patients Receiving DOT

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age, y</th>
<th>Gender</th>
<th>HIV Diagnosis Age, y</th>
<th>CDC Class</th>
<th>Primary Caretaker/ HIV Status</th>
<th>Associated Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>Female</td>
<td>5</td>
<td>A-1</td>
<td>Mother/+</td>
<td>Oral hairy leuokplakia, mother is seriously ill</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>Male</td>
<td>&lt;1</td>
<td>A-1</td>
<td>Mother/+</td>
<td>Chronic supplicative otitis media</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>Female</td>
<td>2</td>
<td>B-3</td>
<td>Great aunt/-</td>
<td>Esophageal candidiasis</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>Female</td>
<td>7</td>
<td>B-2</td>
<td>Mother/+</td>
<td>LIP, maternal mental illness, child unaware of diagnosis</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>Male</td>
<td>&lt;1</td>
<td>B-2</td>
<td>Adoptive mother/-</td>
<td>HIV nephropathy, family disruption</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>Female</td>
<td>1.5</td>
<td>B-3</td>
<td>Grandparents/-</td>
<td>LIP, FTT, gatroscopy</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>Female</td>
<td>&lt;1</td>
<td>B-3</td>
<td>Mother/+</td>
<td>Maternal substance abuse, sister with HIV (patient 9)</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>Female</td>
<td>6</td>
<td>B-2</td>
<td>Mother/+</td>
<td>FTT, chronic parotitis, gatroscopy, family unaware of diagnosis</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>Female</td>
<td>&lt;1</td>
<td>B-3</td>
<td>Mother/+</td>
<td>Maternal substance abuse, sister with HIV (patient 7), depression, gatroscopy</td>
</tr>
</tbody>
</table>

LIP indicates lymphocytic interstitial pneumonitis; FTT, failure to thrive; −, HIV negative; +, HIV positive.

a This patient had 2 DOT admissions.
b This patient had 3 DOT admissions.

the patient was discharged and followed in clinic. This improvement was long-standing in some patients (patient 4 has a response lasting >24 months). The second pattern of response seen (Fig 1B) was similar to the first in the hospital phase, with a marked decrease of VL, but not sustained in the home environment (patient 5, 7, and 9), necessitating reemphasizing the importance of adherence and identifying specific reasons for being nonadherent. The third pattern was seen in patients in whom the VL failed to improve with DOT (Fig 1C). In these patients, hospitalization was helpful in recognizing the limitation of current medications (patient 3), and future outpatient strategies included changing the medications to a different regimen or adding medications to the existing regimen based on resistance testing. In another subgroup of patients (patients 5 and 6), a substantial decrease in VL was found at the beginning of DOT as compared with pre-DOT results (Table 2 and Fig 1B). This result suggested previous nonadherence and was available even before the completion of the DOT hospitalization.

We observed that the mean log₁₀ VL had decreased significantly for the DOT period (day 1 to the last day of DOT; P = .004; mean ± SD decrease: 0.8 ± 0.55 log₁₀ HIV-1 RNA copies per mL; 95% confidence interval: 0.34 to 1.25). No statistically significant changes were observed in analysis of other study periods for VL or CD4⁺ T-cell count and percentage. There was, however, a rise in the mean CD4⁺ percentage for the period of time between day 1 of DOT and the 6-month post-DOT follow-up period (mean ± SD increase: 2.88 ± 4.09 CD4⁺ percentage; 95% confidence interval: −0.54 to 6.29), although the rise did not reach statistical significance (P = .09).
DISCUSSION

We made the decision to hospitalize patients for the duration of 1 week. This duration was chosen to make the DOT period as short and effective as possible. The Pediatric AIDS Clinical Trials Group 381 Study has shown a significant drop in HIV-1 VL in infected adolescents receiving HAART after as short as 1 week of therapy.18 Gigliotti et al14 demonstrated, in a small study of HIV-infected children who were heavily drug experienced, significant drops in VL occurring after 4 to 8 days of DOT. Similar results were reported by Roberts et al9 in another small DOT study involving children. We also considered that 7 days of hospitalization provides enough time for both patient and medical staff to get acquainted with the new environment and the concept of DOT, as well as to address, by a multidisciplinary approach, some of the problems that HIV-infected children and adolescents have. Within this short hospitalization period, we indeed found a statistically significant decline in HIV VL of patients.

Nonadherence to medication is a major problem in the care of pediatric and particularly adolescent patients infected with HIV. Differentiating between nonadherence to medications and resistance of the virus to HAART as the cause of virological and immunologic failure is problematic, especially in heavily drug-experienced patients in whom interpreting results of resistance testing is difficult, and remaining therapeutic options are limited. The complexity of associated social problems, some of them unique to the HIV-infected population (Table 1), makes assessment of adherence even more difficult. Particularly in these scenarios, DOT can provide valuable information.

Our management was different for each of the 3 patterns of response that we have seen. In the majority of patients we observed a rapid decline in VL at the end of the DOT period. In these cases, nonadherence was confirmed. Relaying the results to the patient and caregiver together with emphasis on the good results achieved by the patient when the medications were actually taken is an important part of assuring adherence. In patients who maintained the response after the hospitalization, it is prudent both to continuously reinforce the importance of taking medications and assess at the same time for reemergence of nonadherence. Patients who responded initially but had a relapse in the VL after the hospitalization are complex and challenging case subjects who raise the ongoing issues of nonadherence, viral resistance, or both. Therefore, maximizing adherence through the use of DOT will provide valuable information about the ability of a current HAART regimen to suppress viral replication. Addressing specific issues (frequently encountered while the patient is hospitalized for DOT), such as medication-associated toxicities, body image problems, domestic disruption, and depression, can help convince the patients to take their medications. In our adolescent patients, we intensified the clinic visit schedule and created specific adherence group meetings. The meetings included adherence topics discussions and problem solving combined with social

FIGURE 1
Three typical patterns of HIV VL detected over time among HIV-infected patients receiving DOT. HIV VL was measured as RNA log10 copies per mL. Pre indicates pre-DOT clinic visit, Day 1, first day of DOT hospitalization; Post, last day of DOT hospitalization; Follow-up 1, first post-DOT clinic follow-up; 6 mo, 6-month post-DOT clinic follow-up.
activities and provision of small incentives and electronic reminder devices. In 3 patients, we used a subsequent DOT hospitalization to confirm nonadherence and to address multiple associated problems. Two of our patients in this group were adolescent sisters who had admitted, only after the second DOT, to not taking the medications because of gastrointestinal adverse events. Unfortunately, all of the methods that we used to increase adherence in their case have failed, thus demonstrating the complexity of the situation. A statistically significant sustained decrease in the VL in our study population was not observed, probably because of this subgroup of patients. The last pattern was seen in patients in whom the VL failed to improve with DOT. In these patients, hospitalization was helpful both in ruling out nonadherence as the major factor of virological failure and in recognizing the limitation of the current HAART regimen. Future outpatient strategies included changing the medications to a different regimen or adding medications to the existing regimen based on resistance testing. Thus, all 3 types of patients benefited from DOT. In a subgroup of patients, a substantial decrease in VL was evident at the beginning of DOT as compared with pre-DOT results, confirming the “white-coat adherence.” In this phenomenon, patients are improving their medication-taking behavior in the few days before the appointment with the health care provider. We did not encounter this improvement before routine clinic follow-up, thus emphasizing the importance of inpatient DOT in changing both the medication-taking habits of patients and the routine of doctor-patient encounters.

Hospital-based DOT also had other advantages for patients in providing a multidisciplinary approach toward problems encountered while the patient was in the hospital. For example, body image issues associated with a gastrostomy tube in a young adolescent girl were explored, and after a written contract was created between the patient and caregivers, the gastrostomy tube was removed while the patient maintained her weight and medication taking. Addressing medication-associated adverse effects, discussing sexual behavior, providing nutrition advice, and consulting specialists about various medical problems were also accomplished during hospitalization. We have found that the DOT stay provided a relaxed, unhurried atmosphere in which problems could be discussed and addressed, unlike the often-limited time devoted to patients in a busy clinic.

In 2 of our patients (patients 2 and 9 in the third DOT), we have used hospitalization for a specific medical problem (chronic suppurative otitis media for intravenous antibiotics or evaluation of a chronic cough) as an opportunity for DOT. In these circumstances, with the patient being only mildly ill, expected to stay for 7 days, and able to take the medications on time, the DOT protocol (if needed) could be initiated. Not every hospitalization event is suitable for DOT because of the nature and severity of illness, changes in absorption of medications, or problems with the correct timing of medication administration. However, in a well-chosen setting, using the hospitalization also for DOT can be advantageous. In both of our patients, the VL measured at the end of the hospital stay was significantly lower.

CD4+ T-cell count and percentage are independent predictors of disease progression and mortality in HIV-infected children.15 Response to HAART is slower with regard to CD4+ T-cell counts and percentage as evidenced in previous studies.16,21 We did not observe significant changes in our study population immediately after the DOT period, but there was a trend toward increased CD4+ T-cell percentage from day 1 of DOT to the 6-month follow-up. It is possible that, with a larger cohort of patients, we could demonstrate a significant increase in CD4+ T-cell number and percentage and, therefore, provide a benefit in survival.

Our study is limited by a relatively small sample size, the lack of cost-effectiveness analysis, and by its retrospective design. However, given the high frequency of nonadherence to medications in the pediatric and adolescent HIV-infected population and the importance of adherence for longevity, implementing accurate assessment and improvement of adherence is crucial. We have found short, hospital-based DOT as such. Although DOT is a time-honored method, there are only a handful of reports of hospital-based DOT for HIV-infected children and adolescents. Our report is unique; we have used a tertiary care hospital and a uniform protocol of a 7-day DOT for a population of mainly adolescents. This differs from Parsons et al,15 who used DOT in a chronic care facility with a variable length (24–79 days) of stay, or Gigliotti et al,14 who used a hospital stay of 4 to 8 days or an HIV summer camp in a small group of children. Another report published recently by Roberts et al10 describes a 4-day DOT as part of a stepwise interventional approach consisting of home health nurse referral, DOT, and submission of a neglect report in cases of nonadherence. This report involved a small population of young children and focused mainly on nonadherence to medications by the caregivers as medical neglect. Assessing the effect of DOT on VL in this article is difficult, because the VL measurements were done before the actual hospitalization and at the end of the hospital stay combining, therefore, the “white-coat adherence” and DOT effects.

Hospital-based DOT necessitates preparations before implementation of the protocol. In the pre-DOT clinic visit, conveying to the DOT candidates that the hospital stay is not a punishment but is intended to help with medication taking and quality of life is important. Education of the nursing personnel about their unique duties of supervision of medication taking and strict adherence to timing of administration is crucial. The hospital pharmacy should be contacted to ensure the supply of
the exact medications that the patient is taking at home in a timely manner. For example, a problem may be encountered with hospital formulary of combination pills, such as Trizivir (contains abacavir, lamivudine, and zidovudine). Residents should be educated about the DOT concept and about the complexity of problems associated with the HIV-infected population. Also important is the coordination of the multidisciplinary team for using most effectively the hospital stay for purposes of education and counseling.

We propose that a short, hospital-based DOT be considered in appropriate HIV-infected pediatric and adolescent patients suspected of nonadherence, because DOT seems to be an effective method for the assessment and encouragement of medication taking in complex cases.

ACKNOWLEDGMENTS
We thank Drs Surabhi Vora and Kenneth Alexander at the University of Chicago for critical review of the article.

REFERENCES
Upper-Limb Botulinum Toxin A Injection and Occupational Therapy in Children With Hemiplegic Cerebral Palsy Identified From a Population Register: A Single-Blind, Randomized, Controlled Trial

Remo N. Russo, MBBS, FRACP, FAFRM\textsuperscript{a,b,c}, Maria Crotty, FAFRM, PhD\textsuperscript{a,d}, Michelle D. Miller, PhD, MNutrDiet, BSc\textsuperscript{d}, Sonya Murchland, BAppSci, MHSc\textsuperscript{c}, Peter Flett, MBBS, FRACP, FACRM, FAFRM, MRACMA\textsuperscript{e}, Eric Haan, BMedSc, MBBS, FRACP\textsuperscript{f}

\textsuperscript{a}Flinders University Department of Rehabilitation and Aged Care, Repatriation General Hospital, Daw Park, South Australia, Australia; \textsuperscript{b}Department of Child and Adolescent Development and Rehabilitation and \textsuperscript{c}South Australian Clinical Genetics Service, Women’s and Children’s Hospital, North Adelaide, South Australia, Australia; \textsuperscript{d}Novita Children’s Services, Regency Park, South Australia, Australia; \textsuperscript{e}Flinders Centre for Clinical Change and Health Care Research, Flinders University, Bedford Park, South Australia, Australia; \textsuperscript{f}Royal Hobart Hospital, Hobart, Tasmania, Australia

Financial Disclosure: This study was an investigator-initiated and conducted trial coordinated by the Flinders University Department of Rehabilitation and Aged Care. This study was funded by grants from the Financial Markets Foundation for Children and Allergan Australia Pty Ltd.

ABSTRACT

OBJECTIVES. The purpose of this work was to assess the effect of botulinum toxin A and occupational therapy compared with occupational therapy alone on body structure, activities participation, and self-perception in a sample of children (aged 3–16 years) with hemiplegic cerebral palsy recruited from a statewide register.

PATIENTS AND METHODS. Participants of this single-blind, randomized, controlled trial identified from a population-based cerebral palsy register received either an individually prescribed and localized injection of botulinum toxin A with 4 sessions of occupational therapy over 4 weeks (intervention) or occupational therapy alone (control). Outcomes were assessed from 2 domains of the World Health Organization International Classification of Functioning, Disability, and Health: body structure (Modified Ashworth Scale and Tardieu Scale) and activities participation (Assessment of Motor and Process Skills, Goal Attainment Scaling, Pediatric Evaluation of Disability Inventory, and Pediatric Quality of Life Inventory). Self-perception was also measured.

RESULTS. All of the participants (intervention: \(n = 21\); control: \(n = 22\)) provided data at baseline and 3 and 6 months. Mean age was 8.6 years; 23 were boys and 20 were girls. At 3 months, children allocated to receive the intervention performed significantly better in terms of body structure and activities participation. They reported improvements in self-perception for the global self-worth domain. At 6 months, the differences between the intervention and control groups persisted for the measures of body structure but not for activities participation or self-perception.

CONCLUSION. Botulinum toxin A injection combined with a low-intensity occupational therapy program achieves significant improvements in body structure, activity participation, and self-perception.
CEREBRAL PALSY (CP) is defined as a nonprogressive lesion of the immature brain resulting in disorders of movement and postures. It is the most common physical disability in childhood, with an incidence of 2 to 2.5 per 1000 live births. In ~80% to 90% of children with CP, the motor deficit is spasticity. In hemiplegic CP, almost all children achieve independent walking, but many have significant difficulty in every day functional activities because of involvement of the upper limb with more severe limitation in bimanual fine motor activity and reduced motor ability with increasing age. The effect of CP on self-esteem and body image is unclear, but there is evidence of reduced self-concept for adolescent girls with CP. Reduced self-worth may be an issue among all children with CP; however, the magnitude of the deficit and the effectiveness of treatment programs are unknown.

Botulinum toxin A (BTX-A) injection into the upper limb of children with CP has been studied in several nonrandomized trials with varying degrees of improvement in spasticity, cosmesis, and nonvalidated participation level outcome measures. In randomized trials involving children with hemiplegic CP, BTX-A injection into the upper limb has been shown to reduce muscle tone, improve joint range of motion, and improve some aspects of function. Both Fehlings et al and Lowe et al demonstrated an improvement on the Quality of Upper Extremity Skills Test. However this measure, which is tested under verbal instruction, illustrates what the child can do and not what they actually do in real-life situations. Lowe et al also used Goal Attainment Scaling (GAS) to demonstrate that function could improve, indicating a benefit favoring the intervention group. In both of these trials, therapy for the control group was undertaken, and the control group improved on the outcome measures from baseline, but there was no evaluation of the amount or type of therapy given. Two other studies investigating the effects of BTX-A injection in the upper limb quantified the amount of occupational therapy given. In study, a modest dose of therapy (weekly therapy for 6 weeks) resulted in both intervention and control groups showing functional improvement from baseline. In the other study, intensive therapy (30 minutes of therapy 3 times per week for 6 months) resulted in no changes from baseline for the primary outcome measure in both control and intervention groups, so the effect of therapy remains unclear.

Sufficient evidence to support routine treatment has been lacking. There is a need to include reliable and valid outcomes that measure children’s abilities to carry out necessary activities of daily living and meet specified goals at different levels of functioning to evaluate the impact of treatment.

The purpose of this trial was to assess the effect of an individually prescribed and localized injection of BTX-A and occupational therapy compared with occupational therapy alone on body structure, activities participation, and self-perception in a sample of children with hemiplegic CP.

METHODS

Setting and Participants
Children were recruited (June 2004 to September 2005) from the South Australian Cerebral Palsy Register, which contains details of all of the children with CP (confirmed after physical examination) living in metropolitan, rural, and remote areas of South Australia. The study was approved by the research and ethics committees of the Women’s and Children’s Hospital and the Flinders Medical Centre, South Australia.

Children were eligible to participate if they met the following criteria: diagnosis of hemiplegic CP, aged 3 to 16 years, passive joint range of motion within defined limits (elbow extension to neutral, wrist extension to 30° past neutral with fingers extended, supination of the forearm of 30° past neutral, and thumb extension to neutral), ability to initiate movement of the fingers, and a Modified Ashworth Scale (MAS) spasticity score of ≥2 on 4 at the elbow or wrist. Children were ineligible if they had received an injection of botulinum toxin in the upper limb ≤1 year before the study and in the lower limb ≤6 months before the study. Informed consent was obtained from parents (and children cognitively able to consent).

Random Assignment
Random assignment occurred in blocks of 10. The random assignment schedule and envelopes (concealed, opaque, and foil lined) were prepared by an independent statistician using a computer-generated table of random numbers. The Pharmacy Department at the Repatriation General Hospital maintained the envelopes. The research assistant telephoned the Pharmacy Department to obtain the assignment group, organize the referral for occupational therapy, and, if applicable, referred to the pediatric rehabilitation specialist for scheduling of injection. Allocation was recorded in a logbook locked in a filing cabinet and was not revealed to the research occupational therapists at any time.

Intervention Group
Children allocated to the intervention group received individually prescribed and localized injections of BTX-A into the affected upper limb and weekly occupational therapy for 4 weeks. Children were admitted to the day patient ward at the Women’s and Children’s Hospital or the Flinders Medical Centre for injection of the BTX-A under general anesthesia. A muscle stimulator assisted with localization of the muscles injected, because this technique has been shown to greatly improve the accu-
racy of needle placement.27 The muscle stimulator (electrical low frequency stimulator model ELF-001 from Gorman ProMed Pty, Ltd, Melbourne, Australia) is designed to deliver constant current pulses adjustable from 0 to 20 mA intensity with a fixed 0.3-millisecond duration at a rate of 3 pulses per second.

The maximal dose of BTX-A per muscle according to Russman et al28 was followed; however, all of the muscles across the upper limb were injected if tone was affected (tone [MAS] = 0: the muscle was not injected; 1–1.4/4: half the maximal dose was injected; 2–3/4: the maximal dose was injected). Total injected dose did not exceed 12 U/kg of body weight, to a maximum dose of 300 U of Botox (Allergan, Australia Pty Ltd). The dilution of botulinum toxin used was 100 U of Botox per milliliter of normal saline. Postinjection the children were allowed to leave the hospital once they were medically stable. Weekly 1-hour standardized occupational therapy sessions under the supervision of a pediatric occupational therapist were performed over 4 weeks. The focus of each therapy session was on upper-extremity weight bearing, ball skills, fine motor strengthening (through the use of resistive putty-based activities), and bilateral functional activities (which included activities assisting finger agility and dexterity). Before outcome assessment, participants were instructed to avoid revealing treatment allocation to the research staff.

Control Group
The control group received the standardized occupational therapy program. No placebo injections were performed in the control group because of the requirement for general anesthesia.

Data Collection and Instruments
Measures of body structure included a neurologic assessment, the MAS,25 and the Tardieu Scale.29 All of the measures were performed at baseline, 3 months, and 6 months by a pediatric rehabilitation specialist. Participants also reported whether they felt they were the same, worse, or better since the intervention with respect to function and cosmesis. All of the assessments performed by the pediatric rehabilitation specialist were unblinded. This occurred because, at the time of the study, there was only 1 rehabilitation specialist trained to inject botulinum toxin into the upper limb of children using the technique described, and this was also the only specialist available for data collection. Because the principle outcome measures were functional measures and assessments were blinded, the investigators felt that it was acceptable to proceed with the trial with this limitation.

Data for IQ was taken from the Cerebral Palsy Register and was completed for 41 participants (95%). Testing used a number of standardized tests for 28 participants (65%). Where a standardized test was not used, the assessor was asked to rate the level of intellectual function, and this occurred for 15 children (35%).

The primary outcome measures of activity participation were the Assessment of Motor and Process Skills (AMPS) and the GAS. Both measures were undertaken at baseline, 3 months, and 6 months by a trained assessor blind to treatment allocation.

The AMPS is a reliable and valid tool to measure instrumental (complex or domestic) activities of daily living (ADL).30,31 In using the AMPS measure, key skills and actions that facilitate or hinder performance in ADL at the level of expected achievement are identified.30 The assessment is undertaken in the person’s usual environment32 and is consistent with the person-environment-occupational approach of the revised International Classification of Functioning, Disability, and Health.30 A more detailed account of the specific protocol used in this study is provided elsewhere.9

The GAS33 is a sensitive measure used to assess individual goals after treatment34 and is recommended for children undergoing BTX-A injection.26 Desired outcomes are ranked from –2 (much less than expected outcome) to +2 (much greater than expected outcome). The scores are converted to a T score with 50 (SD: 10) as the expected or average outcome score.35

The Self-Perception Profile for Children36 and the Pictorial Scale of Perceived Competence and Social Acceptance for Young Children37 are valid and reliable37–39 measures used to evaluate children’s self-perception. The Self-Perception Profile for Children is a 36-item scale for children >8 years of age and was designed to evaluate domain-specific judgments of children’s perceived competence in the domains of scholastic competence, social acceptance, athletic competence, physical appearance, and behavioral competence, as well as a global perception of self-worth. The Scale of Perceived Competence and Social Acceptance for Young Children was designed for young children up to age 7 years. Twenty-four items evaluating self-perception are shown in picture form representing 2 ends of a continuum, and the child decides which they are most like. The scoring allows evaluation in 4 domains: cognitive competence, physical competence, peer acceptance, and maternal acceptance.37

Other measures assessed at baseline, 3 months, and 6 months by a blind assessor included the Self-Care Domain of the Pediatric Evaluation of Disability Inventory (PEDI)40 and the Pediatric Quality of Life Inventory (PedsQL 4.0).41 The PEDI was scored from parent interviews. The PedsQL was administered to all of the parents and children >4 years of age. Pain was reported at baseline and at the 3- and 6-month follow-up using a visual analog scale.

Serious adverse events (SAEs; an event that was life threatening, fatal, or resulted in hospitalization or permanent disability) were reported by participants and
their families at all of the follow-up assessments. They also reported any other event or complication that they felt was related to the procedure or to their involvement in the trial.

Sample Size
Published AMPS data for children 3 to 12 years old were used to calculate the sample size required to observe an improvement of 0.5 on the AMPS logit scale.30 The AMPS data report a mean of 1.63 and an SD of 0.58. We estimated that the number of children required to detect a significant difference between the 2 groups was 36 (18 per group; assuming power 80%, \( \alpha = .05 \) and a 1-sided test). To account for attrition, we increased the sample size by 20% to 44 (22 per group).

Statistical Analysis
Analyses were on an intention-to-treat basis using SPSS 11.5 (SPSS Inc, Chicago, IL). Means and medians (95% confidence intervals [CIs]) were calculated according to data distribution. Comparisons between groups for categorical variables were made using the \( \chi^2 \) or Fisher’s exact test. For group comparisons of continuous variables, the independent sample \( t \) test or the Mann-Whitney \( U \) test was used. A significance level of .05 was used throughout.

RESULTS
Study Population
Of the 143 children with hemiplegic CP (aged 3–16 years) identified from the South Australian Cerebral Palsy Register, 108 (76%) were assessed, and 51 (47%) were eligible to participate. Forty-three (84% of eligible children) consented and were randomly assigned (Fig 1). Twenty-two were randomly assigned to the control group and 21 to the intervention group. Follow-up data were obtained for all of the participants at 3 and 6 months.

Baseline Comparisons
The demographic, functional, and quality-of-life characteristics of the study groups were similar at baseline (Table 1). The mean (95% CI) age was 8.6 (7.4–9.8) years. The majority of participants were boys (\( n = 23 \)), had hemiplegia of the right side (\( n = 25 \)), were left-hand dominant (\( n = 25 \)), and were not receiving occupational therapy at the time of the study (\( n = 27 \)). Twenty-one participants (49%) reported pain: 13 mild to moderate and 8 severe. The self-concept domain of athletic competence was significantly different at baseline, favoring the control group.

Dose of BTX-A and Occupational Therapy
The mean (SD) units of Botox per kilogram of body weight injected per child was 8.0 U/kg (2.2) with a
minimum dose of 5.0 U/kg and maximum dose of 11.6 U/kg. The dose of occupational therapy received by each participant was equivalent, with no significant differences between the 2 groups for number of sessions (mean, 95% CI: control: 3.41, 3.39–3.43; intervention: 3.81, 3.80–3.82; \( P = .240 \)), time (minutes) spent in therapy per session (mean, 95% CI: control: 50.22, 49.99–50.45; intervention: 51.38, 51.29–51.47; \( P = .81 \)), or time (minutes) for activities within the sessions, which included weight bearing, ball skills, fine motor strengthening, and bilateral functional activities.

**Activity Participation Measures**

At 3 months, children receiving the botulinum toxin injection achieved greater improvements in the GAS and global self-worth than children receiving occupational therapy alone but were worse off for athletic competence (Table 2). At 6 months, none of these differences persisted. Both groups improved their motor and process skill ability of the AMPS from baseline (Fig 2); however, the differences between the groups were not significant. Similarly, there were no significant differences between the groups in the PEDI or PedsQL.

**Body Structure Measures**

At 3 and 6 months, elbow and wrist tone were significantly improved in the intervention group compared with the control group. Similarly, elbow and wrist spasticity improved significantly in the children who received the intervention at 3 and 6 months compared with the control group.

Subjective evaluation of the effects of treatment on function and cosmesis (esthetics) at 3 and 6 months revealed significant differences at both times for function, favoring the intervention participants. For cosmetic appearance more children in the intervention group improved at 3 months compared with the control group (14 of 21 intervention versus 1 of 22 control; \( P = .185 \)).

At 3 months, fewer children reported pain (\( n = 2 \), intervention group; \( n = 2 \), control group) than at baseline. Both groups had 1 participant with mild-to-moderate pain and 1 each with severe-to-overwhelming pain related to falls affecting the upper limb, which resolved spontaneously. At 6 months, 2 children reported mild-to-moderate pain, 1 in each of the study groups.

---

<table>
<thead>
<tr>
<th>TABLE 1 Baseline Characteristics of the Study Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>Demographic</td>
</tr>
<tr>
<td>Age, mean (95% CI), y</td>
</tr>
<tr>
<td>Gender, male/female, ( n )</td>
</tr>
<tr>
<td>Side of hemiplegia, right/left, ( n )</td>
</tr>
<tr>
<td>Dominant hand, right/left, ( n )</td>
</tr>
<tr>
<td>IQ, ( n )</td>
</tr>
<tr>
<td>Above average, ( &gt;110 )</td>
</tr>
<tr>
<td>Average, ( 90–109 )</td>
</tr>
<tr>
<td>Below average, ( 70–89 )</td>
</tr>
<tr>
<td>Cognitive impairment, ( &lt;70 )</td>
</tr>
<tr>
<td>Body structure</td>
</tr>
<tr>
<td>Tardieu( R2-R1 ), mean degrees (95% CI)</td>
</tr>
<tr>
<td>Wrist</td>
</tr>
<tr>
<td>MAS( E ): Elbow, median (95% CI)</td>
</tr>
<tr>
<td>Wrist, median (95% CI)</td>
</tr>
<tr>
<td>Activity participation</td>
</tr>
<tr>
<td>PEDI, mean (95% CI)*</td>
</tr>
<tr>
<td>PEDsQL, mean (95% CI)*</td>
</tr>
<tr>
<td>Parent</td>
</tr>
<tr>
<td>Child</td>
</tr>
<tr>
<td>AMPS( A ): Motor, mean (95% CI)</td>
</tr>
<tr>
<td>Process, mean (95% CI)</td>
</tr>
<tr>
<td>Self-concept( C ):</td>
</tr>
<tr>
<td>Median (95% CI)</td>
</tr>
<tr>
<td>Athletic competence</td>
</tr>
<tr>
<td>Global self worth</td>
</tr>
</tbody>
</table>

\( * \) Data are from an independent Student’s \( t \) test.
\( ^{\#} \) Data are from a Pearson \( \chi^2 \) test.
\( ^{\ddagger} \) Data are from a Mann-Whitney \( U \) test.
\( ^{\dagger} \) Data include only older children ≥8 years old (\( n = 22 \)).
There were 29 adverse events reported by 20 participants over 6 months. Five SAEs were reported by control participants (2 hospital admissions for seizures in 1 child with epilepsy and 3 hospital admissions for medical reasons in another). There were no minor adverse events reported by the control group. There were 23 adverse events that occurred in the intervention group, 1 of which was an SAE in a child with epilepsy (admission to hospital after a seizure). The most frequently reported adverse events during the study were feeling unwell after the anesthetic (vomiting and cough) in 4 children.

<table>
<thead>
<tr>
<th>TABLE 2: Outcome of the Study Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>AMPSSa, Motor, mean (95% CI)</td>
</tr>
<tr>
<td>3 mo</td>
</tr>
<tr>
<td>6 mo</td>
</tr>
<tr>
<td>Process, mean (95% CI)</td>
</tr>
<tr>
<td>3 mo</td>
</tr>
<tr>
<td>6 mo</td>
</tr>
<tr>
<td>GAS, mean (95% CI)a</td>
</tr>
<tr>
<td>3 mo</td>
</tr>
<tr>
<td>6 mo</td>
</tr>
<tr>
<td>Self-conceptb,c</td>
</tr>
<tr>
<td>Global self worth</td>
</tr>
<tr>
<td>3 mo</td>
</tr>
<tr>
<td>6 mo</td>
</tr>
<tr>
<td>Athletic competence</td>
</tr>
<tr>
<td>3 mo</td>
</tr>
<tr>
<td>6 mo</td>
</tr>
<tr>
<td>PEDI, mean (95% CI)a</td>
</tr>
<tr>
<td>3 mo</td>
</tr>
<tr>
<td>6 mo</td>
</tr>
<tr>
<td>PEDsSQL, mean (95% CI)a</td>
</tr>
<tr>
<td>3 mo</td>
</tr>
<tr>
<td>6 mo</td>
</tr>
<tr>
<td>Tardieu, mean degrees (95% CI)a</td>
</tr>
<tr>
<td>3 mo R2-R1</td>
</tr>
<tr>
<td>Elbow</td>
</tr>
<tr>
<td>Wrist</td>
</tr>
<tr>
<td>6 mo R2-R1</td>
</tr>
<tr>
<td>Elbow</td>
</tr>
<tr>
<td>Wrist</td>
</tr>
<tr>
<td>MAS scores, median (95% CI)a</td>
</tr>
<tr>
<td>3 mo</td>
</tr>
<tr>
<td>Elbow</td>
</tr>
<tr>
<td>Wrist</td>
</tr>
<tr>
<td>6 mo</td>
</tr>
<tr>
<td>Elbow</td>
</tr>
<tr>
<td>Wrist</td>
</tr>
<tr>
<td>Reported functional effectsd</td>
</tr>
<tr>
<td>3 mo</td>
</tr>
<tr>
<td>Worse</td>
</tr>
<tr>
<td>No change</td>
</tr>
<tr>
<td>Better</td>
</tr>
<tr>
<td>6 mo</td>
</tr>
<tr>
<td>Worse</td>
</tr>
<tr>
<td>No change</td>
</tr>
<tr>
<td>Better</td>
</tr>
</tbody>
</table>

---

*a* Data are from an independent Student’s *t* test.

*b* All other domains showed no significant differences.

*c* Data are from a Mann-Whitney *U* test.

*d* Data are from a Pearson *χ²* test.
and excessive weakness in the injected limb in 5 chil-
dren, which was prolonged in 2 of the 5. Headache was
reported by 2 participants. Flu-like symptoms were ex-
perienced by 1 child occurred on a hot day, and he recov-
ered with rest and fluids. He had experienced these
episodes in the past. One adolescent participant experi-
cenced anxiety, and 1 experienced depression, but each
had similar episodes in the past, and both recovered
without specific intervention. One child experience alo-
pelia and had skin scrapings, which confirmed fungal
infection, and this was treated. The family of 1 child with
fatigue felt that this was related to activities at the end of
the school term.

**DISCUSSION**

This single-blind, randomized, controlled clinical trial
provides modest support for the use of upper-limb
BTX-A injection and therapy in children with hemiplegic
CP. When compared with the control group, we found a
statistically significant improvement in body structure at
3 and 6 months, as well as statistically significant im-
provements in activity participation level function on the
GAS and global self-worth at 3 months.

The group that received therapy and botulinum toxin
experienced an increase in self-worth, whereas the ther-
apy alone group experienced a decrease in this outcome
measure. It is possible that the decrease in self-worth in
the therapy-alone group resulted from an increased fo-
cus on achieving functional outcomes, whereas there
was in fact no real shift in spasticity or tone.

The improvements in self-worth are important given
the critical role that self-esteem has in development.42
Children with hemiplegia are more likely to function in
mainstream school environments alongside their typi-
cally developing peers when compared with other CP
types,43 and improvements in self-worth are likely to
influence their socialization, peer relations, and other
areas of functioning. The use of BTX-A has been noted to
enhance self-esteem by diminishing inappropriate motor
responses44; however, these were not evaluated in our
study. A systematic review of studies that have evalu-
ated self-concept in groups of children with CP10 con-
cluded that adolescent girls with CP have a lower self-
estee than girls without disability, but there was
insufficient evidence to conclude that children with CP
in general have lower self-esteem than their nonaffected
peers. Our study has shown that self-esteem can be
positively influenced and should be measured in future
studies evaluating interventions in children with CP.

In light of the findings in improved self-worth, it is
interesting that there were no reported changes in qual-
ity of life. However, quality of life may depend on mul-
tiple factors not influenced by the intervention or may
reflect a response shift with children’s internal standards and values changing over time.45,46

Although we found a significant difference in athletic competence at 3 months, favoring the control group, this did not persist to 6 months when both groups improved. Athletic competence in the control group was significantly better at baseline, which may explain the significant difference at 3 months and is possibly a result of type I error. Alternatively, the tone-reducing effects of the BTX-A may have altered the child’s perceived ability to perform at his or her usual level of athletic competence. This potential effect requires additional investigation to clarify the preinjection counseling of children and their families. For example, in the context of competitive sporting activities, children and their families may elect to alter the timing of injections to minimize the impact of these changes.

The improvement in the GAS outcome measure for the intervention group at 3 months was statistically significant but not at 6 months because of the control group improving between 3 and 6 months. Children in the intervention group reached their desired goals sooner than the control group and then stabilized, a finding that is consistent with previous work.19 The ability for the intervention to allow the child to realize their stated goals more quickly may have widespread positive effects in relation to their sense of achievement and self-perception and is supported by our finding that global self-worth was also significantly better for the intervention group at this time point. Although results of the PEDI did not improve for the intervention group at 3 months, they did for the control group despite not achieving their stated goals as measured by the GAS. This finding is interesting and difficult to explain, but there were no statistically significant differences in the PEDI at either time point.

Our results support the findings of previous studies on the effect of BTX-A on tone and spasticity.14,17,19 The duration of the effect was well beyond the published therapeutic effect of the botulinum toxin, being 12 to 16 weeks of clinically useful relaxation.4 The prolonged duration of effect may relate to the assessment for tone and spasticity being unblinded at 3 and 6 months or that we injected muscles that may not routinely be injected (tone <2 on a MAS), although this method was not adopted by previous investigators, who also reported a prolonged effect.14,19 There may have been an augmenting effect provided by the occupational therapy. Although there is general agreement that the physical modalities must continue after injection of BTX-A to maximize the treatment episode,47,48 the combined effects of BTX-A injection and therapy remain unclear and warrant additional investigation.49 We quantified the amount of therapy received by all of the participants, but comparison to a true control group (ie, those who received no therapy) was not undertaken.

In general, the injection procedure was well tolerated. The 2 participants who had hospital admissions for prolonged seizures (1 intervention and 1 control) had a history of epileptic seizures before inclusion in the study. One subject in the control group had 3 admissions to hospital for unrelated medical reasons, but this child recovered fully on each occasion. The children experiencing the fainting, fatigue, anxiety, and depression had these problems in the past. The child with the alopecia recovered fully. The remaining adverse events were relatively minor and self-limited and are consistent with the known adverse events occurring with general anesthetic50 and BTX-A injection.28,51 Weakness experienced by 2 participants was prolonged. Excessive weakness has been reported previously.14,17 Prolonged atrophy experienced in masseteric hypertrophy after injection of BTX-A <12 months is known to occur,52 but how this relates to the duration of effect of injection in other muscles is unclear. The individualized injection plans used in this clinical trial resulted in the treatment of muscles affected by a lesser degree of spasticity (MAS 1–1+), which may also be a factor.

A high incidence of pain was reported at the initial assessment. Previous studies have reported a similar incidence of pain in children with CP.53–55 Regardless of where the pain originated, it improved over time for both the treatment and control groups. This is an important finding, and additional evaluation of pain and its treatment in this population of children is required.

Limitations of this study include the inability to give placebo, the single-blinded nature of follow-up, and the lack of true control subjects who received no therapy. Children and their families knew their assignment group. This may have led to unintentional disclosure, which may have influenced the results. These methodologic problems in pediatric research are difficult to overcome.56 Moreover, it is argued that drug trials should test medication against standard therapy and not placebo alone.57 Children in this trial were a subgroup of children with hemiplegic CP and, in general, when considering the inclusion criteria for the trial, could be considered less severely affected.

**CONCLUSIONS**

Botulinum toxin injected into the affected upper limb of children with hemiplegic CP and a low-intensity program of occupational therapy achieves significant improvements in body structure, activity participation, and self-perception. This study adds to previous studies investigating the effects of injection of botulinum toxin in the upper limb of children with CP but is unique given the findings related to improvement in self-worth. Given the relatively short-term improvements over the control group, which were sustained to 6 months at a greater (but not statistically significant) value for self-worth, clinicians can better inform children and their families...
about the potential benefits weighed against the inconvenience and cost of providing this treatment. Furthermore, the potential for repeated injections to sustain the significant improvements found at 3 months warrants additional investigation.

ACKNOWLEDGMENTS
This research protocol was investigator led with funding for study personnel conducted with support from the following sources: Financial Markets Foundation for Children and Allergan Australia Pty Ltd.

We thank Cherie Archer, Lynne Giles, Kylie Lange, and Jacqueline Stepien-Hulleman for data collection, entry, and analyses; the occupational therapists at Novita Children Services for providing the occupational therapy program; and the staff of the South Australian Cerebral Palsy Register, the participants, and the families.

REFERENCES
34. McLaren C, Rodger S. Goal attainment scaling: clinical impli-


41. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: Reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Med Care.* 2001;39:800–812


Predicting Pediatric Distress During Radiation Therapy Procedures: The Role of Medical, Psychosocial, and Demographic Factors

James L. Klosky, PhD, Vida L. Tyc, PhD, Xin Tong, MPH, Deo Kumar Srivastava, PhD, Mindy Kronenberg, PhD, Alberto J. de Armenti, MD, Thomas E. Merchant, DO

Divisions of Behavioral Medicine and Anesthesiology and Departments of Radiation Oncology and Biostatistics, St Jude Children’s Research Hospital, Memphis, Tennessee; Department of Pediatrics, University of Tennessee College of Medicine, Memphis, Tennessee

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

ABSTRACT

OBJECTIVES. The purpose of this work was to identify demographic, medical, and psychosocial variables that predict radiation therapy–related distress among pediatric patients with cancer.

PATIENTS AND METHODS. Seventy-nine children between the ages of 2 and 7 years were consecutively enrolled in the study. Radiation therapy–related distress was measured by rates of anesthesia, observed behavioral distress, and heart rate.

RESULTS. Younger age and higher observed behavioral distress predicted the use of anesthesia, higher baseline heart rate predicted lower initial observed behavioral distress, and prone treatment position was associated with increases in both observed behavioral distress and heart rate relative to baseline.

CONCLUSIONS. Modifiable treatment and psychological variables directly relate to pediatric radiation therapy–related distress. Implementation of developmentally appropriate and cost-effective interventions to reduce procedural radiation therapy distress is warranted.
RADIATION THERAPY (RT) IS A CURATIVE FORM OF THERAPY USED TO TREAT A VARIETY OF PEDIATRIC TUMORS. IT CAN BE USED ALONE OR IN CONJUNCTION WITH SURGERY AND CHEMOTHERAPY FOR DISEASE CONTROL AND TO PRESERVE NORMAL TISSUE STRUCTURE AND FUNCTION.\(^1\) RT IS TYPICALLY DELIVERED TO PATIENTS ONCE DAILY, 5 DAYS PER WEEK, OVER A CONTINUOUS COURSE OF TREATMENT TYPICALLY LASTING FROM 5 TO 7 WEEKS. BEFORE TREATMENT, A 30- TO 90-MINUTE PLANNING SESSION (DESCRIBED AS “SIMULATION”) TAKES PLACE TO CONSTRUCT CUSTOMIZED IMMOBILIZATION DEVICES, LOCALIZE RADIOGRAPHICALLY THE REGION TO BE TREATED, POSITION THE PATIENT FOR TREATMENT, AND PERFORM MEASUREMENTS TO SIMULATE THE GEOMETRY OF THE TREATMENT MACHINE (IE, LINEAR ACCELERATOR).

DISTRESS REACTIONS OCCUR FREQUENTLY AMONG PEDIATRIC PATIENTS UNDERGOING RT DESPITE THE NONINVASIVE AND PAINLESS NATURE OF THE TREATMENT.\(^2,3\) DISTRESS REACTIONS MAY OCCUR AS A RESULT OF UNFAMILIARITY WITH THE PROCEDURE AND MEDICAL STAFF, PAINFUL EXPERIENCES WITH PREVIOUS MEDICAL PROCEDURES, SEPARATION FROM PARENTS AND CAREGIVERS, OR FROM THE SIGHTS AND SOUNDS OF THE RT EQUIPMENT.\(^4\) RT REQUIRES THAT PATIENTS BE IMMOBILIZED FOR EXTENDED TIME PERIODS FOR OPTIMAL TREATMENT DELIVERY TO TAKE PLACE. DAILY REPRODUCIBILITY OF PATIENT POSITIONING ALLOWS FOR MORE PRECISE IRRADIATION OF THE TUMOR SITE WITH SUBSEQUENT REDUCTIONS IN HEALTHY TISSUE IRRADIATION AND ACUTE AND CHRONIC ADVERSE EFFECTS.\(^5\) WHEN CHILDREN ARE UNABLE TO MAINTAIN A FIXED AND REPRODUCIBLE POSITION REQUIRED FOR TREATMENT, THE SUCCESS OF RT IS COMPROMISED, AND ANESTHESIA WILL LIKELY BE REQUIRED.

ALTHOUGH SEDATING/ANESTHETIZING CHILDREN FOR RT IS GENERALLY CONSIDERED SAFE, COMPLETION OF RT PROCEDURES WITHOUT PHARMACOLOGICAL INTERVENTION IS PREFERRED, BECAUSE REPEATED SEDATION, HIGH DOSAGES OF SEDATIVES, MULTIPLE DRUG USE, AND GENERAL ANESTHESIA ALL INCREASE THE RISK OF MEDICAL COMPLICATIONS AMONG CHILDREN.\(^6-8\) THESE COMPLICATIONS MAY INCLUDE TRANSIENT HYPOXIA, LARYNGEAL SPASM, AIRWAY OBSTRUCTION, SINUS ARRHYTHMIA, PROLONGED PROFOUND SEDATION, AND RESPIRATORY DEPRESSION.\(^9,10\) IN SOME CASES, LONGER RECOVERY TIMES MAY RESULT IN REDUCTIONS IN AVAILABLE PATIENT QUALITY TIME THROUGHOUT THE COURSE OF TREATMENT. PATIENTS WHO ARE ABLE TO COMPLY WITH RT WITHOUT PHARMACOLOGICAL INTERVENTION AVOID BOTH THE EATING AND DRINKING RESTRICTIONS, AS WELL AS THE PHYSICAL ADVERSE EFFECTS, ASSOCIATED WITH RECEIVING ANESTHESIA. THEREFORE, IN LIGHT OF THESE RISKS, IT IS PREFERABLE TO IMPLEMENT BEHAVIORAL APPROACHES TO MANAGE RT-RELATED DISTRESS AS AN ALTERNATIVE TO PHARMACOLOGIC MANAGEMENT WHENEVER POSSIBLE.

WHEN CHILDREN UNDERGO COMPLEX OR DEMANDING MEDICAL PROCEDURES, A DECISION MUST BE MADE AS TO WHETHER PHARMACOLOGIC INTERVENTION IS WARRANTED FOR ASSISTING THE PATIENT WITH PROCEDURAL COMPLIANCE. IN MAKING THEIR DECISION, THE MEDICAL TEAM WILL UNDOUBTEDLY BE INFLUENCED BY A NUMBER OF DEMOGRAPHIC, MEDICAL, CHILD, AND PARENTAL PSYCHOSOCIAL VARIABLES THAT HAVE BEEN FOUND TO BE ASSOCIATED WITH INCREASES IN PEDIATRIC PROCEDURAL DISTRESS. AMONG INVASIVE PROCEDURES, YOUNGER AGE,\(^11-15\) FEMALE GENDER,\(^14,16-18\) HIGH CHILD ANXIETY,\(^19,20\) HIGH EXPECTATIONS OF DISTRESS,\(^15,21,22\) AND HIGH PARENT ANXIETY\(^22,23\) HAVE BEEN FOUND TO INCREASE CHILDREN’S PROCEDURAL DISTRESS. INTERESTINGLY, REPORTS OF PEDIATRIC DISTRESS PERSIST EVEN WHEN DEEP SEDATION IS USED DURING MEDICAL PROCEDURES. FOR EXAMPLE, 17% OF PEDIATRIC LEUKEMIA PATIENTS UNDERGOING REPEATED SEDATIONS FOR LUMBAR PUNCTURES AND BONE MARROW ASPIRATIONS REPORTED FEAR SPECIFIC TO SEDATION, 65% COMPLAINED ABOUT FASTING BEFORE THE PROCEDURE AND WAITING FOR THE MEDICAL STAFF, WHEREAS 39% COMPLAINED THAT TOO MANY PEOPLE WERE IN THE ROOM DURING THE SEDATION INITIATION.\(^28\)

MUCH LESS IS KNOWN ABOUT FACTORS RELATED TO PEDIATRIC DISTRESS DURING NONINVASIVE PROCEDURES SUCH AS RT. HARNED AND STRAIN\(^29\) FOUND THAT YOUNGER AGE WAS ASSOCIATED WITH INCREASED SEDATIONS DURING MRI PROCEDURES, PARTICULARLY AMONG THOSE <10 YEARS OF AGE. SIMILARLY, BYARS ET AL\(^10\) FOUND THAT YOUNGER AGE, MALE GENDER, AND NEUROLOGIC IMPAIRMENT WERE ASSOCIATED WITH SEDATION FAILURE AMONG PEDIATRIC PATIENTS UNDERGOING MRI PROCEDURES. SPECIFIC TO RT, YOUNGER AGE HAS BEEN FOUND TO BE THE STRONGEST PREDICTOR OF INCREASED ANTICIPATORY DISTRESS AMONG CHILDREN AGED 2 TO 7 YEARS AWAITING THEIR INITIAL RT PROCEDURE.\(^3\)

THE PURPOSE OF THIS STUDY WAS TO IDENTIFY CHILD AND PARENT SOCIOECONOMIC, MEDICAL, AND PSYCHOSOCIAL VARIABLES THAT BEST PREDICT CHILDREN’S DISTRESS DURING RT SIMULATION. TO DATE, NO STUDIES HAVE EMPIRICALLY EXAMINED PREDICTORS OF RT-RELATED DISTRESS AND SEDATION AMONG PEDIATRIC PATIENTS WITH CANCER. FURTHERMORE, THIS NOVEL STUDY EXAMINES BOTH INITIAL PROCEDURAL DISTRESS (WHEN CLINICIANS TYPICALLY MAKE SEDATION DECISIONS) AND TOTAL PROCEDURAL DISTRESS, AS THEY RELATE TO SEDATION, HEART RATE (HR) AND OBSERVED BEHAVIORAL DISTRESS (OBD) OUTCOMES.

METHODS

Participants

Eighty parents and active patients between the ages of 2 and 7 years, who were receiving outpatient RT in the Department of Radiation Oncology at St Jude Children’s Research Hospital, were consecutively enrolled on this institutional review board-approved protocol. Preschool- and early school-aged children were chosen to participate in this study, because they have historically been most frequently anesthetized but have also evidenced successful completion of RT simulation and treatment without pharmacologic intervention. The presented study was part of a larger 2-group randomized clinical trial designed to reduce RT-related distress among pediatric patients with cancer. The details and findings of this larger study are reported elsewhere.\(^31\)

Eligible patients were those who had a primary diag-
nosis of malignancy, aged 2 to 7 years at the time of RT simulation, used English as their primary language, had no previous experience with external beam irradiation, and were functioning at a level at which they could tolerate RT intervention. Accordingly, eligible patients were required to have an Eastern Cooperative Oncology Group (ECOG) performance status rating of 0 (no physical debilitation or functional impairment) to 3 (moderate physical debilitation or functional impairment) as reported by the child’s physician. One-hundred percent of the families approached for this study participated. One participant was excluded from the analysis, because he arrived to the RT clinic already sedated due to an earlier procedure. As a result, analyses were completed on a sample of 79 participants. Table 1 presents the demographic and medical features for the study group, including gender, race, socioeconomic status, age at simulation, diagnosis, functional status, years since diagnosis, RT treatment position, and parent state anxiety.

**Measures**

**Anesthesia**

A child was considered sedated for the RT simulation if any type of pharmacotherapy intervention was delivered at the time of simulation initiation to insure procedural compliance. Examples of sedation include general anesthesia, intravenous (conscious) sedation, or par oral sedation, alone or in combination. Propofol (Diprivan) was the sedation agent in the vast majority of cases, with midazolam (Versed) and lorazepam (Ativan) being administered less frequently.

**OBD**

Based on the Observation Scale of Behavioral Distress (OSBD) of Jay et al., a behavioral checklist modified for use within the RT setting was used to code behavioral distress experienced during the RT simulation. The OSBD checklist is composed of 12 operationally defined behaviors and is similar in content to other behavioral observation scales and methodologies for rating and scoring children’s behavioral distress during invasive medical procedures. Observed distress behaviors have been reliably coded using the OSBD during both invasive and noninvasive pediatric procedures with correlation coefficients being reported in the 0.72 to 0.99 range across raters.

Trained clinical observers independently rated patients’ distress behavior and recorded the frequency of these behaviors during 5-minute intervals over a 10-minute baseline period and simulation procedure. The behavioral ratings were grouped into 4 categories, which included verbal (eg, “I’m scared”), vocal nonlanguage (eg, crying, screaming, or yelling), body movement/physical manifestations of distress (eg, physical resistance, facial grimacing, or hitting/kicking), and a summed total of these 3 domains. Three master’s level psychologists were trained until observations in each category yielded an interrater reliability correlation coefficient of ≥0.80 over 6 consecutive procedures.

**HR**

Often, physiologic arousal has been included as a measure of distress/anxiety in studies examining pediatric invasive and noninvasive medical procedures; however, there is a lack of consensus as to whether HR is a valid, reliable, and sensitive measure of distress. In this study, a Nellcor N-20 Series handheld pulse oximeter recorded the HR of patients every 30 seconds via an oxisensor attached at the patient’s finger. This pulse oximeter yields reliable measurement of HR ranging from 20 to 250 beats per minute ±3.

**State-Trait Anxiety Inventory**

Parents of children rated their own anxiety via completion of the State-Trait Anxiety Inventory. The State-Trait Anxiety Inventory, a standardized inventory with well-documented clinical validity, is designed to measure state and trait anxiety in adults. Only the State Anxiety Scale was used in this study. The State Anxiety Inventory is composed of 20 self-report items and evaluates how the respondent feels “right now at this moment.” The median α coefficient for the State Anxiety Scale is .93, and test-retest correlations range from .31 to .47, reflecting the transitory nature of state anxiety.

***TABLE 1***

Demographic and Medical Features of Study Participants

<table>
<thead>
<tr>
<th>Variables</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>37 (46.8)</td>
</tr>
<tr>
<td>Male</td>
<td>42 (53.2)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>58 (73.4)</td>
</tr>
<tr>
<td>Black</td>
<td>17 (21.5)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4 (5.1)</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
</tr>
<tr>
<td>CNS</td>
<td>53 (67.1)</td>
</tr>
<tr>
<td>Leukemia</td>
<td>10 (12.7)</td>
</tr>
<tr>
<td>Solid tumor</td>
<td>16 (20.2)</td>
</tr>
<tr>
<td>Socioeconomic status</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>31 (39.2)</td>
</tr>
<tr>
<td>Middle</td>
<td>29 (36.7)</td>
</tr>
<tr>
<td>Low</td>
<td>19 (24.1)</td>
</tr>
<tr>
<td>Functional status (ECOG)</td>
<td></td>
</tr>
<tr>
<td>No impairment (0)</td>
<td>64 (81.0)</td>
</tr>
<tr>
<td>Minimal impairment (1)</td>
<td>10 (12.7)</td>
</tr>
<tr>
<td>Mild impairment (2)</td>
<td>5 (6.3)</td>
</tr>
<tr>
<td>Treatment position</td>
<td></td>
</tr>
<tr>
<td>Supine</td>
<td>45 (57.0)</td>
</tr>
<tr>
<td>Prone</td>
<td>26 (32.9)</td>
</tr>
<tr>
<td>Both</td>
<td>8 (10.1)</td>
</tr>
<tr>
<td>Age at simulation, mean ± SD</td>
<td>4.2 ± 1.6</td>
</tr>
<tr>
<td>Time since diagnosis, mean ± SD, y</td>
<td>0.7 ± 0.9</td>
</tr>
<tr>
<td>Parent state anxiety, mean ± SD</td>
<td>41.5 ± 2.2</td>
</tr>
</tbody>
</table>

Variables total N = 79. CNS indicates central nervous system.
Evidence of validity for this scale is shown by its correlation with other widely used measures of anxiety in adults.

Procedure
Parents of eligible children were approached and presented with the details of the study as they initially arrived with their families for the RT simulation in the Radiation Oncology Clinic at St Jude Children’s Research Hospital. At that time, informed consent was obtained according to institutional guidelines. After study enrollment, participating patients were then observed in a clinic examination room via video monitor, and 10 minutes of baseline data (HR and OBD) were collected immediately before the patient’s RT simulation. During this time, participating parents completed the self-report anxiety inventory. After completion of baseline assessment, the child was taken from the clinic examination room to the RT simulation room, which contained a fluoroscopy unit with a couch and radiograph (including detector). At this time, immobilization devices were constructed, and radiographs were obtained in the treatment position localizing the region to be treated.

Logistics of Observations and Measurements
OBD, HR, and parent-report data were collected during baseline and simulation time points. All of the behavioral observations were made via television monitors located outside the clinic examination and RT simulation rooms, whereas physiologic HR data were collected via a pulse oximeter. Data collection was initiated immediately upon the child’s entrance into the targeted medical room (ie, clinic examination room or the RT simulation room). Once in the RT simulation room, children were given a minimum of 15 minutes to voluntarily comply with the RT procedure. If the child was unable to comply by that time, sedation or anesthesia was prescribed. Simulation, HR, and OBD data were collected until the anesthesia agent was administered or until the simulation procedure was successfully completed.

Data Organization and Reduction
Data were organized to reflect 2 phases of RT simulation. The initial phase scores included the first 3 minutes of HR data and the first 5 minutes of OBD during both baseline and simulation phases of RT. Initial response to a novel treatment room and the impending procedural demands faced by the child has traditionally been associated with considerable behavioral distress and anxiety, as well as decision-making, with regard to procedural compliance. As such, this phase of treatment was considered conceptually distinct from the overall treatment, and data from this interval were considered a separate time point in analysis. The total phase scores include cumulative HR and OBD data collected during the entire baseline and simulation procedures. Distress data (initial and total phases) were reduced to OBD and HR per minute to control for differences in time observed in treatment.

Simulation HR data were not collected for 14 of the participants in our sample because of operator error and, in 1 case, parent refusal. As a result, change scores could not be computed for those participants. As such, HR analyses were conducted with data from 65 subjects.

Statistical Approach
General linear regression model was used to assess whether the following factors (age, gender, race, sedation, treatment position, diagnosis, baseline distress total mean, time since diagnosis, ECOG, treatment group, and parent anxiety baseline) were associated with OBD and HR. The above factors were assessed one at a time with each outcome independently. The factors significant at an α level of .10 were put in the final multiple-regression model. The factors significant at an α level of .05 were considered to affect the outcomes (initial OBD and HR and total OBD and HR). Stepwise logistic regression was used to assess whether the following factors (age, gender, race, diagnosis, time from diagnosis, ECOG, treatment position, total HR, total OBD, observed time during simulation, parent state anxiety baseline, and treatment group) were associated with the sedation rate. All of the factors significant at the .10 level were entered into the multiple regression model. The criteria of significance level to remain in the model were Ps < .05 based on the likelihood ratio test. Two-sided P value was reported.

RESULTS
Anesthesia
At simulation, 62% (49 of 79) of patients required pharmacologic intervention to complete the procedure. Of the 49 participants sedated, 92% received propofol-delivered deep sedation, 6% received intravenous moderate sedation, and 2% received light sedation delivered via par oral sedative. Stepwise logistic regression was used to determine the factors that were significantly associated with anesthesia status (occurrence versus nonoccurrence) at simulation. Age at simulation, gender, race, diagnosis, time from diagnosis, ECOG score, treatment position, total simulation HR mean, total simulation OBD mean, parent state anxiety, and treatment group were included in the final model estimate. Eight patients who received total body irradiation and were treated in both prone and supine positions were excluded from these and all other regression analyses. Age and total simulation OBD scores were significantly associated with anesthesia status. With each 1-year increase of age, patients were 0.09 times less likely to be anesthetized (P < .03), whereas those patients with a 1 unit increase of total OBD during simulation were >1000
times more likely to be anesthetized \((P < .02)\). Logistic regression results are reported in Table 2. Finally, children who were anesthetized experienced significantly more observed distress behavior both at baseline \((P < .001)\) and simulation \((P < .0001)\) as compared with nonsedated patients.

**OBD**

Interrater reliability checks for the behavioral observations were conducted for 20% of the procedures. Good interrater reliability for the verbal, vocal nonlanguage, and body movement/physical distress categories was obtained \((r \text{ values ranged from 0.91 to 0.99})\), indicating reliable observations across raters. As each of the OBD categories were significantly and positively correlated with the total OBD scores \((r \text{ values range from 0.22 to 0.86}; P \text{ values range from } <.06 \text{ to } <.0001)\), only the total OBD scores were used in the analyses.

Means and SDs of OBD scores (initial phase and total) obtained at baseline and simulation, as well as mean change scores (simulation − baseline) for the group, are provided in Table 3. A general linear regression model was used to determine the effect of selected covariates associated with total mean change scores \((P < .02)\). Patients treated in the prone position exhibited a significantly greater increase in the total OBD change scores \((F_{6,64} = 5.26; P < .001)\). Results of this regression analysis are reported in Table 4.

**HR**

Means and SDs for HR outcomes (initial and total phases) for the control and intervention groups with the corresponding mean change scores (simulation − baseline) are shown in Table 3. Simulation HR data and change scores were available for 65 subjects.

Age, gender, race, diagnosis, time since diagnosis, ECOG scores, sedation status, treatment position, mean baseline OBD, and parent state anxiety scores, as well as treatment group, were examined as covariates associated with HR change score outcomes. Time observed in the simulation session was added as a covariate to the model for total HR mean change scores only. Results from the general linear regression model revealed that the treatment position was significantly associated with increases in HR from the initial phase baseline to initial phase simulation \((P < .04)\). Patients treated in the prone position had a significantly greater increase in the initial-phase change scores of mean HR during the initial phase as compared with those treated supine. The final model, including treatment position and sedation status, accounted for 23% of the variance in the initial-phase HR change score outcome \((F_{5,54} = 5.28; P < .003)\).

The final model predicting changes in total mean HR from baseline to simulation included age at simulation, treatment group, treatment position, total mean of baseline OBD, time observed during simulation, and sedation status as independent variables. Treatment position was the only significant covariate associated with total mean HR changes. Those treated in the supine position experienced significantly smaller changes in total mean scores of HR as compared with those treated prone \((P < .02)\). Procedural and behavioral variables accounted for 49% of the variance \((F_{6,51} = 8.18; P < .001)\). Regression results for this model are reported in Table 5.

**DISCUSSION**

The results of our study found that 62% of participants between the ages of 2 and 7 years were anesthetized for RT simulation. In addition, younger children and those with higher levels of behavioral distress were more likely to be anesthetized during simulation. These findings are consistent with reports by our RT staff that OBD, in
combination with age, typically determines the need for sedation, although the level of behavioral distress exhibited by the child is usually subjectively determined. Specific, standardized criteria based on age and defined levels of distress may help to better identify patients that may require anesthesia to complete RT procedures. Furthermore, the association between OBD and HR was found to significantly reduce RT-related distress.4,31 Klosky et al,31 for example, report on a randomized clinical trial that evidenced the efficacy of a cognitive-behavioral intervention in reducing RT-related distress similar to that of orally administered diazepam (Valium) among preschool- and school-aged children.36 In a smaller study, Slifer2 reports that 82% of pediatric patients between the ages of 2 and 7 years were able to complete RT simulation and treatment successfully without any use of sedation after completing a behavioral intervention. Because developmental trends in our sample suggest that rates of sedation are primarily a function of younger age and increased OBD, it is also important that the design and implementation of behavioral interventions be developmentally appropriate and individualized to meet each child’s specific needs. For example, auditory-based distraction/relaxation using age appropriate language and themes should be most useful in reducing RT-related distress among younger patients with visual deficits, whereas visually based modeling or distraction should be most useful for pediatric patients experiencing auditory deficits.

Although the results were generally similar between the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the initial phase and total RT simulation, some differences did emerge.

Another option for reducing RT-related distress includes the use of behavioral and cognitive-behavioral interventions. Techniques such as modeling, distraction, desensitization, positive reinforcement, relaxation, visual imagery, practice, education, and so forth have been found to significantly reduce RT-related distress.4,31 Klosky et al,31 for example, report on a randomized clinical trial that evidenced the efficacy of a cognitive-behavioral intervention in reducing RT-related distress similar to that of orally administered diazepam (Valium) among preschool- and school-aged children.36 In a smaller study, Slifer2 reports that 82% of pediatric patients between the ages of 2 and 7 years were able to complete RT simulation and treatment successfully without any use of sedation after completing a behavioral intervention. Because developmental trends in our sample suggest that rates of sedation are primarily a function of younger age and increased OBD, it is also important that the design and implementation of behavioral interventions be developmentally appropriate and individualized to meet each child’s specific needs. For example, auditory-based distraction/relaxation using age appropriate language and themes should be most useful in reducing RT-related distress among younger patients with visual deficits, whereas visually based modeling or distraction should be most useful for pediatric patients experiencing auditory deficits.

Although the results were generally similar between the initial phase and total RT simulation, some differences did emerge. Specifically, we found that increased initial baseline HR predicted lower initial OBD in the simulation room. It may be that a select group of children experienced anticipatory procedural anxiety as discussed in Blount et al11 and Tyc et al.3 Traditionally, changes in RT-related OBD begin as the simulation procedural demands increase (ie, parent leaves the room.
child positioning is attempted, alone with unfamiliar medical staff, etc), and this type of distress may have been best captured by the total, as opposed to the initial, OBD score. It may be advantageous to design interventions to take place both before and during RT simulation for the most meaningful reductions in RT-related distress to occur.

Although the results of this study represent a positive step in identifying factors that contribute to children’s RT-related distress, the findings should be interpreted in the context of the study’s limitations. Sedation histories and number and type of previous medical procedures experienced by the patients in our sample were not assessed, although these variables may influence the child’s distress and response to the RT procedure. Likewise, we did not assess children’s perceptions of their own distress because of the young age of the study participants and their inability to provide valid self-reports. True natural baseline HR and behavioral data (ie, at home or in their natural setting the day before the RT simulation) were also not obtained, making it difficult to examine “base rate” behavior and HR levels that could potentially influence our results. Initial distress may have also been overrepresented in the study, because total procedural distress scores included initial distress as well. The results discussed here should not be generalized beyond the parameters of the study: children 2 to 7 years of age undergoing RT simulation.

The results of our study indicate that both fixed and modifiable variables directly relate to distress as experienced by pediatric patients with cancer undergoing RT simulation. Developmentally appropriate interventions designed to target these variables among preschool- and early school-aged children are clearly warranted. Furthermore, incorporation of empirically tested strategies to improve child procedural coping (ie, filmed modeling, distraction, education, etc) are needed to maximize successful outcomes as evidenced by reductions in distress and rates of sedation. Future studies examining the efficacy of such proposed interventions are needed and should be tested in both the RT-simulation and treatment settings.

ACKNOWLEDGMENTS

This study was funded in part by a grant from the Starlight Starbright Foundation, the American Lebanese Syrian Associated Charities, and the National Cancer Institute (Cancer Center Support grant CA21765).

We thank nursing, radiation therapy, and anesthesia personnel for their assistance with this study.

REFERENCES

32. Hollingshead AB. *Four Factor Index of Social Status.* New Haven, CT: Yale University; 1975
Exposure to Movie Smoking Among US Adolescents Aged 10 to 14 Years: A Population Estimate

James D. Sargent, MD, Susanne E. Tanski, MD, Jennifer Gibson, MS

Norris Cotton Cancer Center, Dartmouth Medical School, Lebanon, New Hampshire

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

ABSTRACT

BACKGROUND. Several studies have linked seeing smoking in movies with adolescent smoking, but none have determined how much movie smoking adolescents see.

OBJECTIVE. Our aim was to determine exposure to movie smoking in a representative sample of young US adolescents.

METHODS. We surveyed 6522 nationally representative US adolescents aged 10–14 years. We content analyzed 534 contemporary box-office hits for movie smoking. Each movie was assigned to a random subsample of adolescents (mean: 613) who were asked whether they had seen the movie. Using survey weights, we estimated the total number of US adolescents who had seen each movie and then multiplied by the number of smoking depictions in each movie to obtain gross smoking impressions seen by adolescents.

RESULTS. The 534 movies were mainly rated PG-13 (41%) and R (40%), and 74% contained smoking (3830 total smoking occurrences). On average, each movie was seen by 25% of the adolescents surveyed. Viewership was higher with increased age and lower for R-rated movies. Overall, these movies delivered 13.9 billion gross smoking impressions, an average of 665 to each US adolescent aged 10–14 years. Although this sample’s R-rated movies contained 60% of smoking occurrences, they delivered only 39% of smoking impressions because of lower adolescent viewership. Thirty popular movies each delivered ≥100 million gross smoking impressions. Thirty actors each delivered >50 million smoking impressions, such that just 1.5% of actors delivered one quarter of all character smoking to the adolescent sample.

CONCLUSIONS. Popular movies deliver billions of smoking images and character smoking depictions to young US adolescents. Removing smoking from youth-rated films would substantially reduce exposure from new box-office hits. Furthermore, the popular actors who frequently smoke in movies could have a major impact on adolescent movie smoking exposure by choosing not to portray characters who smoke.
REPORTS OF AN association between seeing smoking in movies and adolescent smoking\textsuperscript{6–7} have prompted greater scrutiny of smoking depicted in movies. Movie smoking has been quantified through content analyses, in which a specified sample frame of movies is assessed according to a set of predefined criteria. The published content analyses agree on multiple points: smoking is depicted in most movies\textsuperscript{6–14}; movie smoking increases as the Motion Picture Association of America (MPAA) rating increases from G to R\textsuperscript{8,13,14}; movie smoking rarely is associated with negative health outcomes\textsuperscript{8,9,11,14}; and movie characters who smoke are more affluent than the typical US smoker.\textsuperscript{8,11} Studies that report the prevalence of character smoking concur that 20\% to 25\% of characters in popular contemporary movies smoke\textsuperscript{12,15,16} and that smoking rates are higher among male characters. One area that has been less thoroughly assessed is how much of the smoking in movies actually reaches adolescents. Movie smoking seen by a relatively small proportion of adolescents would be unlikely to have a large population effect on smoking. For movies popular among adolescents, however, a single smoking depiction may reach a large audience. In this context, as it is in the advertising arena, an assessment of the number of smoking depictions delivered to a defined audience is considered a key metric of the success of the campaign.

One 2004 non–peer-reviewed publication estimated the number of smoking impressions delivered to US adolescents.\textsuperscript{17} By using an Internet parental review/screening service (www.screenit.com) for some 800 box-office releases to estimate smoking content and Nielsen figures to calculate the youth audience, they determined that the selected movies delivered some 8.3 billion smoking scenes to children and teens, or \textasciitilde1350 per capita. Despite a higher concentration of smoking in R-rated movies, youth-rated movies delivered approximately half of the smoking scenes. The limitations of the study notwithstanding, this report underlines the potential impact of movie smoking because of the extensive reach into the adolescent population. A recent article looking at images of smoking included in 31 television trailers for movies found that even these brief advertisements resulted in substantial exposure, with 270 million smoking impressions delivered to youth in a single year.\textsuperscript{18} Our aim in this study is to assess exposure to movie smoking in US adolescents aged 10 to 14 years using more rigorous quantitative methods that allow for an evaluation of smoking within a popular contemporary sample of films and aggregated at the levels of movie and leading actors.

METHODS

Between June and October, 2003, we conducted a random-digit dial telephone survey of 6522 US adolescents aged 10 to 14 years. We have previously reported details of the survey, including the survey completion rate (66\%).\textsuperscript{1} Briefly, the telephone surveys were conducted by trained interviewers using a computer-assisted telephone interview system, and interviewers were trained to administer the survey in English or Spanish. We obtained parental consent and adolescent assent before interviewing each respondent. To protect confidentiality, adolescents indicated their answers to sensitive questions by pressing numbers on the telephone, rather than speaking aloud. All aspects of the survey were approved by the institutional review boards at Dartmouth Medical School and Westat. The sample was broadly representative of the US adolescent population (as assessed by the US Census) with respect to age, gender, household income, and census region (Appendix 1). Compared with the 2000 US Census, our unweighted sample had higher percentages of Hispanics and slightly lower percentages of blacks; we adjusted for these differences using a post-stratification weighting procedure.

One aim of the survey was to assess the proportion of adolescents who had seen each movie of a sample of popular contemporary movies. We selected the top 100 US box-office hits per year for each of the 5 years preceding the survey (1998–2002; \(N = 500\)) and 34 movies that earned at least $15 million in gross US box-office revenues during the first 4 months of 2003. The computer-assisted telephone interview survey was programmed to randomly select 50 movie titles from the larger pool for each adolescent interview, such that adolescents were randomly assigned to movie titles, and a representative subsample of adolescents (mean number surveyed: 613; SD: 27) was queried on each title. Movie selection was stratified by the MPAA rating so that the distribution of movies in each list reflected the distribution of the full sample of movies (19\% G/PG, 41\% PG-13, and 40\% R). Respondents were asked whether they had ever seen each movie title on their unique list.

To assess the possibility of false-positive responses, we asked all of the adolescents whether they had seen a sham movie title, \textit{Handsome Jack}, and <2\% reported having seen it. We have shown previously that adolescents reliably remember movies they have seen 1 to 2 years before a survey.\textsuperscript{2}

Content Analysis

Trained coders counted the number of smoking occurrences in each of the 534 movies using previously validated methods.\textsuperscript{3} A smoking occurrence was counted whenever a major or minor character handled or used tobacco in a scene (defined as a “smoking episode”) or when background smoking was present (defined as a smoking “incident”). The smoking occurrence was used as the unit of movie exposure in characterizing adolescent movie smoking exposure by movie. To determine exposure to movie smoking according to actor, we restricted the analysis to the smoking episodes by major
and minor characters. A major character was defined by the movie coders as someone “central to the development of the plot”; coders identified a median of 6 (interquartile range: 5–8) major characters in each movie. Smoking episodes were counted irrespective of the scene’s duration or how many times the tobacco product appeared during the scene. We used 2 movie coders and double-coded 10% of the movies. As a measure of interrater agreement, the Pearson correlation coefficient for the number of smoking episodes (character smoking) contained in each double-coded movie was 0.99 and for smoking incidents (background smoking or signage) was 0.86.

Defining Exposure
In assessing advertising campaigns, it is important to determine how many individuals or households are exposed to a particular advertising media or message; this is what marketing researchers refer to as a determination of reach. Because adolescents were randomly assigned to a movie, we assumed that the proportion of adolescents within each subsample who had seen each movie reflected the proportion of adolescents in the US population who had seen the movie, hereafter referred to as the “movie viewership.” For each movie, we multiplied the movie viewership proportion by the number of US adolescents aged 10 to 14 years in 2002 (20.88 million) to obtain an estimate of the number of US adolescents who had seen the movie, our measure of reach. “Gross impressions” is the total number of exposures delivered by a media schedule. For movies, gross smoking impressions were determined by multiplying the estimate of the number of US adolescents who had seen the movie by the number of smoking occurrences in the movie. Per capita gross impressions of movie smoking were obtained by dividing the total number of gross smoking impressions across all of the movies in the sample by the total US population of adolescents. This measure is similar to that used to determine the success of a marketing campaign, the gross rating point, except that our measure is conservative in that it does not account for multiple viewing of 1 movie by an adolescent. In our determination of gross smoking impressions delivered by leading actors, we aggregated by actor, the movie smoking occurrences delivered by all of the major characters in ≥1 movie.

RESULTS
As reported previously,1 the movie sample was composed mainly of PG-13 (41%) and R-rated (40%) movies (Table 1) and contained a total of 3830 smoking occurrences. Sixty percent of all smoking occurrences appeared in R-rated movies, 36% in PG-13 movies, 4% in PG movies, and <1% in G-rated movies. Mean movie viewership was 0.25 (SD: 0.21). Figure 1 shows the distribution for movie viewership by MPAA rating for the 534 movies. The mean of movie viewership was similar for G-rated movies (0.51) and PG-rated movies (0.47) but significantly lower for PG-13 movies (0.27) and for R-rated movies (0.14; P < .001). Figure 2 illustrates the distribution of movie viewership by age of the respondent. Movie viewership increased in a linear fashion with age (P < .001 for trend), from a mean of 0.19 for 10-year-old adolescents to 0.34 for 14-year-olds.

Gross Smoking Impressions Delivered Overall
In aggregate, the 534 movies delivered some 13.9 billion impressions of smoking to this age group (Table 1). Examined on a per capita basis, an average of 665 impressions of smoking were delivered to each US adolescent aged 10 to 14 years as of the date of the survey, September 2003. Despite containing a higher number of smoking occurrences, R-rated movies accounted for a lower proportion of gross smoking impressions delivered to US adolescents (39%) than PG-13 movies (50%) because of lower viewership of R-rated movies. Figure 3 illustrates this: youth-rated movies contained ~40% of smoking occurrences, yet they delivered ~60% of all gross smoking impressions to young adolescents.

Gross Smoking Impressions Delivered According to Movie
Table 2 lists all of the movies that delivered ≥100 million gross smoking impressions (N = 30). The movies on the list shared 2 common characteristics: they included many smoking impressions and were popular, seen by 2 to 15 million of the estimated 21 million US 10- to 14-year-olds. Not all of the movies were intended for youths in this age range. For example, the R-rated movie Hannibal contained graphic violence scenes, yet was seen by ~25% of adolescents in this age range. Most of the movies on this list depicted cigarette smoking. However, Lord of the Rings: The Fellowship of the Ring depicted pipe smoking, Wild Wild West depicted cigar smoking, and Atlantis: The Lost Empire depicted animated cigarette smoking.
Gross Smoking Impressions According to Major and Minor Characters

In this movie sample, we identified 7720 major or minor movie characters (Table 3). The smoking rate among minor characters was much lower than for major characters. Among the 3813 major characters, 71% were male (we did not code gender for minor characters), and the smoking rate was slightly higher among males. Movie characters accounted for 2760 (72.1%) of the 3830 smoking occurrences in the movies, delivering 9.88 billion gross impressions of character smoking to US adolescents. The majority of the character gross smoking impressions (65.2%) were delivered by male major characters.

Of the 1961 actors who played character roles in these films, 499 (25.4%) smoked in ≥1 role. Table 4 shows the number of smoking characters, total smoking episodes, and gross smoking impressions for 30 actors who delivered ≥50 million gross impressions of character smoking to US adolescents (see Appendix 2 for details on movies and names of smoking characters). These 30 actors accounted for 2.4 billion gross smoking impressions, or 24.6% of all of the gross smoking impressions delivered by characters in these movies. Of the 30 actors shown, only 3 were women, in keeping with the observation that most actor smoking is done by males. Generally,
high impact actors played smoking characters in several movies released over the study period; but some actors, such as Keanu Reeves, delivered all of their smoking in a single popular movie (Hardball). Some of the characters delivering the high numbers of smoking impressions were not human: Florence Stanley delivered the highest number of smoking impressions of any actor by playing the voice for Mrs Packard, the chain-smoking communications officer on the submarine in the animated movie Atlantis: The Lost Empire, and Ian McKellan played Gandolf, a pipe-smoking wizard in The Lord of the Rings.

As mentioned above, the majority of actors did not smoke in films. Moreover, a number of actors starred in multiple films in this sample and did not smoke in any of them. Table 5 lists all of the actors who starred in ≥5 films during the study period and portrayed no smoking characters. Table 5 includes the names of many notable actors, some of whom (eg, Tom Cruise) have rarely, if ever, smoked in movies, and others (eg, Bruce Willis), who were known previously for playing smoking characters in highly popular movies (eg, Die Hard).

**DISCUSSION**

This is the first study to assess the delivery of movie smoking impressions to US adolescents using survey techniques that allow for a direct estimate of the exposure. The results indicate that this sample of popular contemporary movies delivered billions of smoking im-
pressions to American youth and provide a basis, from a communications standpoint, for the large population effect (an adjusted attributable risk of 0.50 in 1 longitudinal study) of movie smoking on adolescent smoking. A single popular movie with smoking can deliver tens of millions of smoking depictions to adolescents on the first run at the box office. Once that film also appears on DVD, video, and movie channels, the movie extends its reach and may ultimately deliver hundreds of millions of gross smoking impressions to youth. These findings represent a conservative estimate of the impact of movies, because we did not assess exposure to all of the films, did not determine how many times adolescents had seen each film, and restricted our survey to a both a narrow age range and date range for the movie sample. Indeed, children begin watching animated movies with smoking as preschoolers, and, with the penetration of VCRs and DVD players, are able to view their favorite films over and over throughout childhood and adolescence. Older movies are also readily available, resulting in the delivery of many more smoking depictions from these films than we were able to document in this study.

When examined according to movie rating, we found that, although R-rated movies accounted for the majority of occurrences of movie smoking, they delivered only ~40% of smoking impressions to adolescents. This assessment corresponds closely with the findings by Po-

### Table 4
Character Smoking for Actors Who Delivered ≥50 Million Gross Smoking Impressions to US Adolescents Aged 10 to 14 Years

<table>
<thead>
<tr>
<th>Actor Name</th>
<th>Character Portrayals</th>
<th>No. of Smoking Episodes</th>
<th>Gross Smoking Impressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>-----------------</td>
<td>----------------------</td>
<td>-------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Stanley, Florence</td>
<td>0 1 15</td>
<td>183.3</td>
<td></td>
</tr>
<tr>
<td>Hawkes, John</td>
<td>1 3 18</td>
<td>156.8</td>
<td></td>
</tr>
<tr>
<td>Avari, Erick</td>
<td>2 2 11</td>
<td>147.4</td>
<td></td>
</tr>
<tr>
<td>McKellan, Ian</td>
<td>2 1 10</td>
<td>125.0</td>
<td></td>
</tr>
<tr>
<td>De Niro, Robert</td>
<td>7 4 28</td>
<td>117.9</td>
<td></td>
</tr>
<tr>
<td>Gibson, Mel</td>
<td>4 3 21</td>
<td>90.9</td>
<td></td>
</tr>
<tr>
<td>Borgnine, Emer</td>
<td>0 1 9</td>
<td>89.8</td>
<td></td>
</tr>
<tr>
<td>Reeves, Keau</td>
<td>4 1 10</td>
<td>84.7</td>
<td></td>
</tr>
<tr>
<td>Pitt, Brad</td>
<td>3 3 42</td>
<td>83.7</td>
<td></td>
</tr>
<tr>
<td>Dahn, Wemer</td>
<td>0 1 9</td>
<td>83.5</td>
<td></td>
</tr>
<tr>
<td>Close, Glenn</td>
<td>1 1 5</td>
<td>82.0</td>
<td></td>
</tr>
<tr>
<td>Branagh, Kenneth</td>
<td>1 2 6</td>
<td>77.7</td>
<td></td>
</tr>
<tr>
<td>Gould, Elliott</td>
<td>0 1 12</td>
<td>76.6</td>
<td></td>
</tr>
<tr>
<td>Reilly, John C</td>
<td>5 4 12</td>
<td>75.7</td>
<td></td>
</tr>
<tr>
<td>Hackman, Gene</td>
<td>5 4 36</td>
<td>73.9</td>
<td></td>
</tr>
<tr>
<td>Fichtner, William</td>
<td>5 1 8</td>
<td>70.1</td>
<td></td>
</tr>
<tr>
<td>McDonald, Christopher</td>
<td>3 1 5</td>
<td>69.6</td>
<td></td>
</tr>
<tr>
<td>Glover, Crispin</td>
<td>2 1 5</td>
<td>66.3</td>
<td></td>
</tr>
<tr>
<td>Jackman, Hugh</td>
<td>1 3 11</td>
<td>63.4</td>
<td></td>
</tr>
<tr>
<td>Goldberg, Adam</td>
<td>2 3 13</td>
<td>61.9</td>
<td></td>
</tr>
<tr>
<td>Ifans, Rhy</td>
<td>1 2 16</td>
<td>61.2</td>
<td></td>
</tr>
<tr>
<td>Voight, Jon</td>
<td>4 3 8</td>
<td>59.3</td>
<td></td>
</tr>
<tr>
<td>Kline, Kevin</td>
<td>1 3 5</td>
<td>57.4</td>
<td></td>
</tr>
<tr>
<td>Mortensen, Viggo</td>
<td>3 2 5</td>
<td>56.7</td>
<td></td>
</tr>
<tr>
<td>Carrey, Jim</td>
<td>3 2 9</td>
<td>55.7</td>
<td></td>
</tr>
<tr>
<td>McBride, Chi</td>
<td>2 3 8</td>
<td>53.8</td>
<td></td>
</tr>
<tr>
<td>Reedus, Norman</td>
<td>0 1 8</td>
<td>53.4</td>
<td></td>
</tr>
<tr>
<td>Washington, Denzel</td>
<td>5 3 11</td>
<td>52.9</td>
<td></td>
</tr>
<tr>
<td>Hopkins, Anthony</td>
<td>6 3 14</td>
<td>52.7</td>
<td></td>
</tr>
<tr>
<td>Cage, Nicolas</td>
<td>2 7 37</td>
<td>50.4</td>
<td></td>
</tr>
</tbody>
</table>

*Data are shown in millions.

### Table 5
Characters Who Starred in ≥5 Films But Did Not Smoke

<table>
<thead>
<tr>
<th>Actor Name</th>
<th>Character Portrayals</th>
<th>No. of Smoking Episodes</th>
<th>Gross Smoking Impressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson, Anthony</td>
<td>12 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affleck, Ben</td>
<td>11 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Willis, Bruce</td>
<td>10 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cruise, Tom</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cummings, Jim</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>David, Keith</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freeman, Morgan</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hanks, Tom</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isaacs, Jason</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee, Jason</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moore, Julianne</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shalhoub, Tony</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shannon, Molly</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spacey, Kevin</td>
<td>7 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baker, Dylan</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clooney, George</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gooding, Cuba Jr</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lawrence, Martin</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Levy, Eugene</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lindo, Delroy</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lopez, Jennifer</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miller, Larry</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nealon, Kevin</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prince, Freddie Jr</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sioce, Kath</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomas, Sean Patrick</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Union, Gabrielle</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ward, Fred</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Williams, Robin</td>
<td>6 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balaban, Bob</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bates, Kathy</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boomer, Mika</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleece, John</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curtis, Cliff</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detmer, Amanda</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dick, Andy</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duval, Robert</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ferrell, Will</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kennedy, Jamie</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Levine, Ted</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liotta, Ray</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mohr, Jay</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morton, Joe</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myers, Mike</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neeson, Liam</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patton, Will</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paxton, Bill</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piven, Jeremy</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapaport, Michael</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suplee, Ethan</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sutherland, Donald</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woodard, Alfre</td>
<td>5 0 0 0 0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Data are shown in millions.
Y
t tilted view of movie smoking. We would expect movie smoking to vary in its influence on adolescent smoking, depending on the individual characteristics of the adolescent and the context of the movie smoking depiction. At the movie level, our counts of movie smoking include character and background depictions of smoking and do not differentiate on the duration of the occurrence. In addition, our character smoking measure did not distinguish between different types of smoking (cigar, pipe, or cigarette) or other contextual elements, such as the type or attractiveness of character who smokes, whether the smoking occurred in an emotional or arousing scene, or the overall level of violence in the movie. One would expect, for example, that cigarette smoking depicted by Rusty Ryan (Brad Pitt) in Oceans Eleven might have more impact on adolescent smoking than pipe smoking by Gandolf (Ian McKellan) in Lord of the Rings. In addition, we do not consider differential impact by adolescent gender or race/ethnicity. More research needs to be done on whether these factors impact adolescent smoking behavior differentially.

Despite its limitations, this study offers the first direct assessment of the reach of movie smoking, providing a measure of the magnitude by which this medium delivers potentially powerful social influence impressions to a vulnerable population. The finding that movies deliver smoking depictions by the billions to US adolescents during a period when they are susceptible to social influences to smoke warrants further research to better understand the process by which these depictions might affect behavior. Of additional importance, with more than half of box-office receipts for US movies coming from overseas, US movies deliver smoking impressions to adolescents all over the world. We hope that this research will prompt those in the movie industry who make decisions about movie smoking to think carefully about the role they may be playing in the smoking epidemic both domestically and worldwide. Finally, we encourage pediatric practitioners to support the efforts of the American Academy of Pediatrics and other groups to pressure the movie industry to adopt voluntary incentives and other policies to limit smoking in movies. In addition, pediatricians should counsel parents regarding the media, monitoring steps that they can take to reduce movie smoking exposure to their children. As stated in the conclusion of a recently published article on the topic, “parental rules and monitoring of children’s movie viewing may have a protective influence on children’s risk for smoking and drinking. over and above parental monitoring of nonmedia-related behaviors.”

ACKNOWLEDGMENTS

We acknowledge funding from the National Cancer Institute.

We thank our collaborators, Madeline Dalton, Mi-
Michael Beach, Jennifer Tickle, and Todd Heatherton, and also the content coders, Daniel Nassau, Balvinder Rakhra, and Elaini Bergamini, for their efforts.

REFERENCES


APPENDIX 1 Demographic Characteristics for 6522 Adolescents Aged 10 to 14 Years in the Baseline Survey Compared With US 2000 Census

<table>
<thead>
<tr>
<th>Data</th>
<th>Baseline Sample %</th>
<th>US Census %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unweighted</td>
<td>Weighted</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>11</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>13</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>14</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>Female</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>66</td>
<td>65</td>
</tr>
<tr>
<td>Black</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>American Indian</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>81</td>
<td>84</td>
</tr>
<tr>
<td>Census region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Midwest</td>
<td>24</td>
<td>22</td>
</tr>
<tr>
<td>South</td>
<td>34</td>
<td>36</td>
</tr>
<tr>
<td>West</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Household income, $1000s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>20–29</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>30–50</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>50–75</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td>&gt;75</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>Actor Name</td>
<td>Movie</td>
<td>Smoking Character</td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Stanley, Florence</td>
<td>Atlantis: The Lost Empire</td>
<td>Mrs Packard</td>
</tr>
<tr>
<td>Hawkes, John</td>
<td>I Still Know What You Did Last Summer</td>
<td>Dave</td>
</tr>
<tr>
<td></td>
<td>The Perfect Storm</td>
<td>Michael “Bugsy” Moran</td>
</tr>
<tr>
<td></td>
<td>Hardball</td>
<td>Ticky Tobin</td>
</tr>
<tr>
<td>Avari, Erick</td>
<td>Mr Deeds</td>
<td>Cecil Anderson</td>
</tr>
<tr>
<td></td>
<td>Master of Disguise, The</td>
<td>Cigar Maker</td>
</tr>
<tr>
<td>McKellan, Ian</td>
<td>Lord of the Rings: The Fellowship of the Ring, The</td>
<td>Gandalf</td>
</tr>
<tr>
<td>De Niro, Robert</td>
<td>Ronin</td>
<td>Sam</td>
</tr>
<tr>
<td></td>
<td>Adventures of Rocky &amp; Bullwinkle, The</td>
<td>Fearless Leader</td>
</tr>
<tr>
<td></td>
<td>Men of Honor</td>
<td>Billy Sunday</td>
</tr>
<tr>
<td></td>
<td>15 Minutes</td>
<td>Eddie Flemming</td>
</tr>
<tr>
<td>Gibson, Mel</td>
<td>What Women Want</td>
<td>Nick Marshall</td>
</tr>
<tr>
<td></td>
<td>Payback</td>
<td>Porter</td>
</tr>
<tr>
<td></td>
<td>We Were Soldiers</td>
<td>Lt Col Hal Moore</td>
</tr>
<tr>
<td>Borgnine, Ernest</td>
<td>Small Soldiers</td>
<td>Kip Killigan</td>
</tr>
<tr>
<td>Reeves, Keanu</td>
<td>Hardball</td>
<td>Conor O'Neill</td>
</tr>
<tr>
<td>Pitt, Brad</td>
<td>Fight Club</td>
<td>Tyler Durden</td>
</tr>
<tr>
<td></td>
<td>Snatch</td>
<td>Mickey O'Neill</td>
</tr>
<tr>
<td></td>
<td>Ocean’s Eleven</td>
<td>Rusty Ryan</td>
</tr>
<tr>
<td>Dahm, Werner</td>
<td>XXX</td>
<td>Kirill</td>
</tr>
<tr>
<td>Close, Glenn</td>
<td>102 Dalmatians</td>
<td>Cruella De Vil</td>
</tr>
<tr>
<td>Branagh, Kenneth</td>
<td>Wild Wild West</td>
<td>Dr Arliss Loveless</td>
</tr>
<tr>
<td></td>
<td>Road to El Dorado, The</td>
<td>Miguel</td>
</tr>
<tr>
<td>Gould, Elliott</td>
<td>Ocean’s Eleven</td>
<td>Reuben Tishkoff</td>
</tr>
<tr>
<td>Reilly, John C</td>
<td>Thin Red Line, The</td>
<td>Sgt Storm</td>
</tr>
<tr>
<td></td>
<td>Perfect Storm, The</td>
<td>Dale “Murph” Murphy</td>
</tr>
<tr>
<td></td>
<td>Chicago</td>
<td>Amos Hart</td>
</tr>
<tr>
<td></td>
<td>Hours, The</td>
<td>Dan Brown</td>
</tr>
<tr>
<td>Hackman, Gene</td>
<td>Behind Enemy Lines</td>
<td>Reigart</td>
</tr>
<tr>
<td></td>
<td>Replacements, The</td>
<td>Jimmy McGinty</td>
</tr>
<tr>
<td></td>
<td>Royal Tenenbaums, The</td>
<td>Royal Tenenbaum</td>
</tr>
<tr>
<td></td>
<td>Heartbreakers</td>
<td>William B. Tentsy</td>
</tr>
<tr>
<td>Fichtner, William</td>
<td>Perfect Storm, The</td>
<td>David “Sully” Sullivan</td>
</tr>
<tr>
<td>McDonald, Christopher</td>
<td>Iron Giant, The</td>
<td>Kent Mansley</td>
</tr>
<tr>
<td>Glover, Crispin</td>
<td>Charlie’s Angels</td>
<td>Thin Man</td>
</tr>
<tr>
<td>Jackman, Hugh</td>
<td>X-Men</td>
<td>Logan/Wolverine</td>
</tr>
<tr>
<td></td>
<td>Swordfish</td>
<td>Stanley</td>
</tr>
<tr>
<td></td>
<td>Someone Like You…</td>
<td>Eddie Alden</td>
</tr>
<tr>
<td>Goldberg, Adam</td>
<td>Saving Private Ryan</td>
<td>Pvt Mellish</td>
</tr>
<tr>
<td></td>
<td>EDTV</td>
<td>John</td>
</tr>
<tr>
<td>Ilfans, Rhys</td>
<td>Beautiful Mind, A</td>
<td>Sol</td>
</tr>
<tr>
<td></td>
<td>Notting Hill</td>
<td>Spike</td>
</tr>
<tr>
<td></td>
<td>Replacements, The</td>
<td>Nigel Gruff</td>
</tr>
<tr>
<td>Voight, John</td>
<td>Pearl Harbor</td>
<td>President Roosevelt</td>
</tr>
<tr>
<td></td>
<td>Ali</td>
<td>Howard Cosell</td>
</tr>
<tr>
<td></td>
<td>Holes</td>
<td>Mr Sir</td>
</tr>
<tr>
<td>Kline, Kevin</td>
<td>Wild Wild West</td>
<td>Artemus Roosevelt</td>
</tr>
<tr>
<td></td>
<td>Wild Wild West</td>
<td>President Grant</td>
</tr>
<tr>
<td></td>
<td>Road to El Dorado, The</td>
<td>Tulio</td>
</tr>
<tr>
<td></td>
<td>Orange County</td>
<td>Marcus Skinner</td>
</tr>
<tr>
<td>Mortensen, Viggo</td>
<td>Lord of the Rings: The Fellowship of the Ring, The</td>
<td>Aragorn</td>
</tr>
<tr>
<td>Carrey, Jim</td>
<td>Man on the Moon</td>
<td>Tony Clifton 1</td>
</tr>
<tr>
<td></td>
<td>Me, Myself &amp; Irene</td>
<td>Charlie</td>
</tr>
<tr>
<td></td>
<td>Me, Myself &amp; Irene</td>
<td>Hank</td>
</tr>
<tr>
<td>McBride, Chi</td>
<td>Gone in Sixty Seconds</td>
<td>Donny Astricky</td>
</tr>
<tr>
<td></td>
<td>Undercover Brother</td>
<td>The Chief</td>
</tr>
<tr>
<td></td>
<td>Cradle 2 the Grave</td>
<td>Jump Chambers</td>
</tr>
<tr>
<td>Redus, Norman</td>
<td>Blade II</td>
<td>Scut</td>
</tr>
<tr>
<td>Washington, Denzel</td>
<td>Fallen</td>
<td>John Hobbes</td>
</tr>
<tr>
<td></td>
<td>Siege, The</td>
<td>Anthony Hubbard</td>
</tr>
<tr>
<td></td>
<td>Training Day</td>
<td>Alonzo</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Actor Name</th>
<th>Movie</th>
<th>Smoking Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hopkins, Anthony</td>
<td><em>Mask of Zorro, The</em></td>
<td>Zorro/Don Diegode la Vega</td>
</tr>
<tr>
<td></td>
<td><em>Hannibal</em></td>
<td>Hannibal Lecter</td>
</tr>
<tr>
<td></td>
<td><em>Hearts in Atlantis</em></td>
<td>Ted Brautigan</td>
</tr>
<tr>
<td>Cage, Nicolas</td>
<td><em>Snake Eyes</em></td>
<td>Rick Santoro</td>
</tr>
<tr>
<td></td>
<td><em>City of Angels</em></td>
<td>Seth</td>
</tr>
<tr>
<td></td>
<td><em>8mm</em></td>
<td>Tom Welles</td>
</tr>
<tr>
<td></td>
<td><em>Bringing out the Dead</em></td>
<td>Frank Pierce</td>
</tr>
<tr>
<td></td>
<td><em>Windtalkers</em></td>
<td>Joe Enders</td>
</tr>
<tr>
<td></td>
<td><em>Family Man, The</em></td>
<td>Jack Campbell</td>
</tr>
<tr>
<td></td>
<td><em>Captain Corelli’s Mandolin</em></td>
<td>Captain Antonio Corelli</td>
</tr>
</tbody>
</table>
Altering Portion Sizes and Eating Rate to Attenuate Gorging During a Fast Food Meal: Effects on Energy Intake

Cara B. Ebbeling, PhD, Erica Garcia-Lago, BA, Michael M. Leidig, RD, Linda G. Seger-Shippee, DT, Henry A. Feldman, PhD, David S. Ludwig, MD, PhD

Division of Endocrinology, Department of Medicine, and Clinical Research Program, Children’s Hospital Boston, Boston, Massachusetts

ABSTRACT

OBJECTIVE. Eating large amounts of food at a rapid rate, defined as gorging, may contribute to excess energy intake. We aimed to evaluate whether altering portion sizes and eating rate could decrease energy intake during an extra-large fast food meal.

METHODS. Subjects were adolescents (n = 18), 13 to 17 years of age, who reported eating fast food ≥1 time per week. BMI exceeded the 80th percentile for all subjects. Three feeding conditions were evaluated with a crossover design. Total amounts and types of foods and beverage served during the meal were held constant across conditions, equaling ~125% of that consumed during a baseline assessment visit when subjects were offered unlimited amounts. The meal (chicken nuggets, French fries, and cola) was presented as 1 large serving at a single time point (condition A, standard), portioned into 4 smaller servings presented at a single time point (condition B, effects of portioning), or portioned into 4 smaller servings presented at 15-minute intervals (condition C, effects of portioning and eating rate). Energy intake across conditions was compared by using analysis of variance.

RESULTS. Energy intake was not significantly different, whether expressed in kilojoules (mean ± SEM: condition A, 5552 ± 357 kJ; condition B, 5321 ± 433 kJ; condition C, 5762 ± 500 kJ) or relative to total daily energy expenditure (mean ± SEM: condition A, 51.9 ± 3.5%; condition B, 48.2 ± 4.0%; condition C, 53.0 ± 4.3%).

CONCLUSIONS. Adolescents consumed ~50% of energy needs regardless of manipulations in portion sizes and eating rate to attenuate gorging. This finding suggests that nutritional factors inherent to fast food, such as low levels of dietary fiber, high palatability, high energy density, high fat content, high glycemic load, and high content of sugar in liquid form promote excess energy intake.
Fast food is ubiquitous, and frequent consumption may be among the high-impact dietary behaviors that contribute to excess weight gain in adolescents. An estimated 75% of adolescents eat fast food ≥1 times per week,1 representing a dramatic increase since the 1970s.2 Escalating portion sizes of menu items3 and increasing frequency of fast food meals4 have paralleled the increasing prevalence of obesity.4 Parallel trends raise the possibility of a causal relationship between fast food consumption and the obesity epidemic. Moreover, several studies have shown a direct association between fast food consumption and body weight or energy intake.5-7

Many characteristics of fast food meals have been shown to promote energy intake in feeding studies, including enormous portion sizes,8 rapid eating rate,10 low levels of dietary fiber,11 high palatability,12 high energy density,13 high fat content,14 high glycemic load,15 and high content of sugar in liquid form.16 Of these, only 2, namely, enormous portion sizes and rapid eating rate, are not nutritional factors inherent to fast food. The verb “to gorge” means “to swallow in large mouthfuls or quantities” or “with greediness,”17 which describes consumption of enormous portion sizes at a rapid rate.

Strong evidence links portion sizes with energy intake.8,9,18,19 Large portion sizes may override internal homeostatic mechanisms that regulate satiety, possibly by distorting visual cues and decreasing awareness of food consumption.19,20 In a study using “bottomless bowls” to alter portion sizes inconspicuously, Wansink et al20 concluded that individuals eat more when they are unable to monitor intake visually. Rolls et al19 came to a similar conclusion on the basis of a study showing increased energy intake with progressively larger portions of snacks provided in opaque packages.

The effects of large portion sizes on energy intake also may be attributable, in part, to a rapid eating rate.8,10 Kral et al10 found a direct correlation between eating rate and meal size when obese adults were offered unlimited portions of a liquid test meal. Fisher et al9 attributed increased energy intake to accelerated eating rate, as indicated by an increase in bite size with no change in bite frequency, when young children were served relatively large versus age-appropriate portions of a lunch entrée. A rapid eating rate may not allow adequate time for development of physiologic satiety signals involved in meal termination, although data are inconsistent.10,21-24

We showed previously that excess energy intake is a characteristic outcome when adolescents, particularly those who are overweight, are presented with an extra-large fast food meal.25 Conceptually, an extra-large meal is a stimulus that fosters the behavioral response of gorging, with the outcome being excess energy intake. The purpose of this study was to determine whether reducing portion sizes and slowing eating rate, to attenuate gorging, would decrease energy intake during a fast food meal. This study differs from previous studies of portion sizes8,9,18,19 in that we evaluated portioning (ie, dividing an extra-large meal into smaller fractions, without altering absolute amounts of foods and beverage), as opposed to portion control (ie, serving smaller absolute amounts). If visual cues enhance awareness of consumption and thereby curb energy intake, then simple measures involving portioning and packaging of fast food might prove useful, from a public health perspective, in preventing and treating adolescent obesity. If visual cues do not enhance awareness, then fundamental improvements in nutritional factors inherent to fast food, a prospect that has been historically challenging,26 would seem to be warranted in efforts aimed at combating the obesity epidemic.

METHODS

Subjects
We enrolled 20 adolescents, 13 to 17 years of age, who reported eating fast food ≥1 time per week. Screening for inclusion and exclusion criteria was conducted through telephone interviews with the adolescent and a parent. Adolescents with BMI values exceeding gender- and age-specific 85th percentile values,27 on the basis of self-reported weight and height data, were invited to participate in the study. Weight and height were measured at the initial visit to ensure that all enrolled subjects had BMI values exceeding the 80th percentile, allowing for some inaccuracy in self-reported data. We excluded adolescents who reported having a diagnosis of a major medical illness or eating disorder, smoking ≥1 cigarette in the past week, or taking any prescription medication that may affect food intake. During the telephone interviews, we collected demographic data, including gender, ethnicity and race, and date of birth.

The protocol was approved by the institutional review board at Children’s Hospital Boston (trial registration NCT00121706 [see www.clinicaltrials.gov]). Newspaper advertisements and fliers, stating that the purpose of the study was to collect information on how teenagers eat fast food, were used to recruit subjects. At the time of recruitment, we confirmed that the subjects did not have aversions to menu items in the fast food meal served during study visits (ie, chicken nuggets, French fries, and cola). We did not mention strategies for altering portion sizes and eating rate. Written informed consent and assent were obtained from parents and subjects, respectively. Data were collected during the summer of 2005. As incentive, we offered each subject $150 in gift certificates, with $25 provided at each of 4 visits and an additional $50 provided after completion of 6 telephone-administered, 24-hour dietary and physical activity recall interviews.
Overview
The initial study visit was considered a baseline assessment visit for establishment of individualized amounts of foods and beverage that would be served during each of 3 subsequent test visits. We instructed subjects to eat a standard breakfast of cold cereal and milk at 9:00 AM on the day of each visit and then not to eat or to drink anything, except water, until after the visit. At 1:30 PM, we served a meal from a national fast food chain in the naturalistic setting of a food court. The length of each meal was 1 hour. Meals were served to groups of 3 subjects, on average, to foster socializing, which is characteristic of the fast food experience for adolescents. Boys and girls were in separate groups, to avoid any effects of social interactions between the genders that might influence eating behaviors.

Before each meal, we asked each subject to rate his or her level of hunger by using a 10-cm visual analog scale, anchored with the descriptors “not at all hungry” and “extremely hungry.” After the meals, we asked each subject to rate the meal size, relative to the size of fast food meals that he or she consumed typically, by using a 10-cm visual analog scale ranging from “much smaller than usual” to “much larger than usual.”

Weight and height were measured at each visit by using an electronic scale (model BWB-800; Tanita, Tokyo, Japan) and stadiometer (model PE-AIM-101; Perspective Enterprises, Portage, MI), respectively. We calculated BMI as weight in kilograms divided by the square of height in meters.

Baseline Assessment Visit
During visit 1, we evaluated energy intake in an extra-large meal containing conventional fast food (Table 1), by using methods modeled after our previous study.25 The following standard instructions were read to the group of subjects before the meal: “We will bring each of you a meal. Eat as much or as little as you like, until you have had enough. There is more food available, and you may eat as much as you want. Please do not share your food with others in the group. If you need more of anything, just ask. Keep your packaging on your tray.” Research staff members monitored food intake discreetly. Whenever approximately three fourths of the extra-large meal portion of chicken nuggets or fries had been consumed, a refill portion of the item was added to the tray (Table 1). Empty cola containers were replaced with full containers immediately. Ketchup and sweet and sour sauce could be obtained from the center of the table throughout the meal. By using this standardized protocol, we provided more of the items that each subject enjoyed most and would be likely to order in large portions when given the option.

Energy intake was calculated on the basis of the difference in weight between the amount of each menu item served and that of leftovers. In preparation for data collection, we purchased and weighed 20 reference units for each menu item. The mean weights of the reference units were used to estimate the amounts of menu items served (Table 1). Leftovers were weighed to the nearest 1 g on an electronic digital scale (model TLC-100; Tanita Corp.). Nutrition Data System for Research (NDSR) software (version 2005; Nutrition Coordinating Center, University of Minnesota, Minneapolis, MN) was used to convert the weight of each item consumed to energy intake.

Test Visits
By using a crossover design for visits 2 to 4, we assigned each subject randomly to 1 of 6 possible sequences of 3 feeding conditions. The random assignment was stratified according to gender. Identification numbers for male participants were matched randomly to a single block of 12 assignments (ie, with each possible feeding sequence represented twice) and those for female participants to 2 blocks of 12 and 6 assignments. The assignments were prepared on index cards by the study statistician and were delivered in opaque envelopes to the principal investigator, to be opened after each participant’s baseline assessment visit.

Meal delivery varied according to condition, as described in Table 2, with the total amounts of foods and beverage held constant across conditions. In brief, the fast food meal was presented as 1 large serving at a single time point (condition A, standard), portioned into 4 smaller servings presented at a single time point (condition B, effects of portioning), or portioned into 4 smaller servings presented at 15-minute intervals (condition C.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Fast Food Meal Fed During Baseline Assessment Visit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Menu Item</strong></td>
<td><strong>Extra-Large Meal</strong></td>
</tr>
<tr>
<td>Chicken nuggets</td>
<td>10 pieces, 160 g</td>
</tr>
<tr>
<td>French fries</td>
<td>1 large serving, 172 g</td>
</tr>
<tr>
<td>Cola*</td>
<td>1 bottle, 20 fl oz</td>
</tr>
<tr>
<td>Ketchup</td>
<td>4 packets, 32 g</td>
</tr>
<tr>
<td>Sweet and sour sauce</td>
<td>2 packets, 60 g</td>
</tr>
</tbody>
</table>

1 kcal = 4.2 kJ. Energy values represent data derived from the NDSR and are based on the mean weights of the reference units. The total energy value for the extra-large meal was 5691 kJ.

* A refrigerated bottle of cola, rather than a cup of cola, was provided to avoid measurement inaccuracies associated with variable amounts of ice.
TABLE 2  Feeding Conditions for Test Visits

<table>
<thead>
<tr>
<th>Condition</th>
<th>Meal Delivery and Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Standard</td>
<td>The total amount of food was delivered on a tray at time 0. Instructions were as follows: “We will bring you a meal of chicken nuggets, French fries, and cola. Eat as much or as little as you like, until you have had enough. Please do not share your food with others in the group. Keep your packaging on your tray.”</td>
</tr>
<tr>
<td>B. Effects of portioning</td>
<td>The total amount of food was divided equally among a tray and 3 lunch boxes, with the tray and all of the boxes being delivered at time 0. Instructions were as follows: “We will bring you 4 servings of chicken nuggets and French fries, 1 on your tray and 3 in separate boxes. Each serving contains exactly the same foods. Begin with what is on your tray. Eat as much or as little as you like. When you’re through with the first box, you may then open and eat from the next box. You do not have to finish all of the foods in one box before going on to the next box. There are also extra cups of cola. When you are finished with the cup on your tray, you may take another cup. Continue with one box after another, and one cup after another, until you have had enough. Please do not share your food with others in the group.”</td>
</tr>
<tr>
<td>C. Effects of portioning and eating rate</td>
<td>The total amount of food was divided equally among a tray and 3 lunch boxes, with the tray being delivered at time 0 and the boxes being delivered at regular intervals (15 min, 30 min, and 45 min). Instructions were as follows: “We will bring you a meal of chicken nuggets, French fries, and cola. Eat as much or as little as you like. Three additional servings will be delivered to you in separate boxes during the lunch. Each serving contains exactly the same foods. We will also deliver additional servings of cola. When each additional box arrives, you may then open and eat from it. You do not have to finish all of the foods in one box before going on to the next box. Continue with one box after another, until you have had enough. Please do not share your food with others in the group.”</td>
</tr>
</tbody>
</table>

Effects of portioning and eating rate). The total amount of each item other than condiments in the meal (ie, nuggets, fries, and cola), whether presented as a single large serving or portioned into 4 smaller servings, was equal to ~125% of that consumed during the baseline assessment. We did not provide any refill portions during the test visits.

After purchasing the nuggets and French fries from the fast food establishment, we immediately portioned and served these items for the conditions outlined above. Foods were removed from the vendor packaging and served in small, medium, or large containers so that subjects would have appropriate visual cues regarding portion sizes. Similarly, chilled cola was transferred from large bottles to small, medium, or large cups. Ketchup (4 packets per serving) and sweet and sour sauce (2 packets per serving) also were provided, with refill portions readily available in the center of the table throughout the meal.

Energy intake, the primary study outcome, was calculated on the basis of the difference in weight between the amount of each menu item served and that remaining on the tray or in boxes at the end of the meal. All containers were weighed to the nearest 1 g on an electronic digital scale before and after the meal, and NDSR was used to convert the weight of each item consumed to energy intake, as described above.

Dietary and Physical Activity Recall Interviews

We collected dietary and physical activity data during telephone-administered, 24-hour recall interviews, by calling each subject on the 2 days after each of the 3 test visits to assess behaviors during the day of the visit and the day after the visit. The focus of this report is on data corresponding to the day of each visit, with particular attention directed toward recall of fast food consumed during the visit. The interviewer was masked with respect to the sequences of conditions.

Dietary intake was assessed with a multiple-pass method using NDSR. The interviewer prompted the subject to list in sequence the foods and beverages consumed during the previous day, to identify omissions in the initial list, and then to provide details (eg, portion sizes and brand names) concerning each reported item. Intake was reviewed and confirmed at the end of each interview. Recalled energy intake during the fast food meal at the food court provided a measure of awareness with regard to fast food consumption.

To assess physical activity, we asked the subject to recall the activity performed most during respective 15-minute time blocks throughout the preceding day (12:00 AM to 11:59 PM) and then to rate the relative intensity as light, moderate, hard, or very hard.28 This protocol also provided information regarding inactivity (eg, sleeping or television viewing). A metabolic equivalent level was assigned to each activity for calculation of a physical activity factor. As points of reference, resting has a metabolic equivalent level of 1.0 and brisk walking has a level of 5.0.29 Total energy expenditure was estimated by multiplying the basal metabolic rate, calculated with validated equations,30 by the physical activity factor.

Before the interviews, we held in-person group training sessions on how to estimate food and beverage portion sizes and how to describe the intensity of physical activity. The training sessions occurred after the fast food meal during the baseline assessment visit. Subjects practiced recalling dietary intake and physical activity during the training sessions.
**Statistical Methods**

We hypothesized that energy intake would be higher when the fast food meal was presented as 1 large serving at a single time point (condition A), compared with multiple smaller servings at 15-minute intervals (condition C). In addition, we hypothesized that energy intake would be intermediate when the meal was presented as multiple servings at a single time point (condition B). To test this hypothesis, analyses were conducted by using SAS 9.0 statistical software (SAS Institute, Cary, NC).

We compared study outcomes across the 3 feeding conditions by using repeated-measures analysis of variance. All results are presented as mean ± SEM. Statistical significance was defined as \( P < .05 \). The analysis of variance included a fixed effect to test for systematic variation across the 3 successive visits (order effects) and an interaction term to test whether differences among feeding conditions depended on the position in the sequence (effect modification). To account for correlated outcomes within subject, we used a compound-symmetry covariance structure for the repeated-measures analysis, equivalent to a random subject effect in ordinary analysis of variance.

Power analysis for the 3-period crossover design was based on an estimate of intersubject variability in energy intake\(^{31}\) and a conservative estimate of 0.2 for intrasubject, interday correlation. The sample of 18 provided 80% power to detect a difference as small as 689 kJ between condition A and condition C, our primary comparison, with the specified statistical criterion (\( P < .05 \)). For the 2 secondary comparisons (condition A versus condition B and condition B versus condition C), we used \( P < .025 \) as a critical level, following the Bonferroni rule, and had 80% power to detect a difference of 769 kJ.

**RESULTS**

Eighteen of the 20 subjects (4 male subjects and 14 female subjects) enrolled in the study completed all of the study visits. Descriptive characteristics are presented in Table 3. One of the subjects was Hispanic or Latino (race not reported). Among the others, the racial distribution was as follows: 3 white, 11 black, 3 not reported. Subjects consumed a mean of 5809 ± 441 kJ at the baseline assessment visit, when instructed to eat as much or as little food as desired. The hunger and meal size ratings were 5.5 ± 0.6 cm and 7.6 ± 0.5 cm, respectively.

Study outcomes are presented in Table 4. There were no differences across conditions for energy intake during the fast food meal, whether expressed in kilojoules or relative to total energy expenditure. Position in the visit sequence had no systematic effect on intake \( (P > .29) \), and there was no significant interaction between feeding condition and visit number \( (P > .12) \).

Of note, none of the subjects consumed all of the foods and beverage provided; consequently, energy served exceeded energy intake for all subjects under all conditions (Table 4). The gram weights of foods and beverage consumed and ratings of hunger and meal size also did not differ across conditions. Nevertheless, subjects recalled energy intake more accurately for condition C, compared with condition A. Physical activity and energy expenditure during the day of the feeding visit did not differ across conditions.

**DISCUSSION**

It could be argued that the very business model of the fast food industry promotes gorging. Advertising campaigns frequently emphasize the extra value of enormous portion sizes offered for an apparently minimal increase in cost over smaller orders.\(^{26}\) Free beverage refills in some establishments exemplify this strategy. Moreover, the industry was founded on the premise of quick service, as conveyed by the term “fast food,” and relies on rapid customer turnover to secure profits.\(^{32,33}\)

Large portion sizes\(^{8,9,18,19}\) and a rapid eating rate\(^{8,10}\) were shown previously to promote excess energy intake under various conditions. Moreover, strategies aimed at portion control consistently curb energy intake, presumably because of lack of available food. In contrast to previous studies,\(^{8,9,18,19}\) our aim was not to evaluate whether adolescents consume less when given smaller absolute amounts of foods and beverage but rather to explore the effects of portioning and eating rate while eliminating confounding by amounts provided. In the present study, energy intake during a fast food meal was not influenced by portioning (condition A versus condition B) or eating rate (condition B versus condition C). Assuming a dietary pattern of 3 meals and 1 or 2 snacks per day, the average meal size to achieve energy balance would be ~25% to ~30% of total energy needs, equating to ~2730 to ~3275 kJ on the basis of the estimated energy expenditure for subjects in the present study. Given that subjects consumed ~5460 kJ under all conditions, they overate regardless of manipulations in food presentation. Subjects recalled fast food intake more accurately after condition C, compared with condition A, presumably because portioning and slowing the eating rate, in combination, provide visual cues that enhance awareness of intake. Nevertheless, meal size ratings did

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean ± SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>153 ± 0.3</td>
</tr>
<tr>
<td>Body weight, kg</td>
<td>78.6 ± 4.3</td>
</tr>
<tr>
<td>Height, cm</td>
<td>163.5 ± 7.5</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>29.2 ± 1.3</td>
</tr>
<tr>
<td>BMI percentile(^a)</td>
<td>95.9 ± 1.4</td>
</tr>
</tbody>
</table>

\(^a\) BMI percentile values were calculated on the basis of gender- and age-specific growth curves.\(^{37}\)
not differ across conditions, and any increase in awareness clearly had no effect on total energy intake.

Our findings do not simply reflect the “clean plate phenomenon” (ie, consuming all food provided, regardless of amounts), because all subjects had leftovers on their trays or in their lunch boxes at the end of every fast food meal. Therefore, subjects might have eaten to their physical limits at all test visits. To evaluate this possibility, we estimated the volumes corresponding to the gram weights of consumed foods and beverage, assuming that the average density after chewing and swallowing is ~1 g/mL. The total volume consumed, averaged across all 3 conditions, was 1074 mL. Although our calculations do not account for gastric emptying during the meals, this figure compares to the maximal capacity of the stomach (~800–1000 mL) based on maximal ratings of abdominal discomfort in gastric distention studies of obese individuals.34,35 Moreover, energy intake during the test visits was similar to what we observed during the baseline assessment visit, when food was replaced continuously to ensure that subjects never had a clean plate.

The propensity of overweight individuals to consume large amounts is not restricted to fast food meals. However, eating foods with low energy density (such as fruits, vegetables, legumes, and minimally processed grain products), even to physical limits, would not promote excess energy intake to the same extent as consuming conventional fast food.36 The energy density of fast food is extremely high, relative to prevailing dietary patterns,37 because of its very low content of water and fiber and very high content of fat, starch, and added sugar. Moreover, chemical manipulations often are used to achieve high palatability,33 which also may promote excess energy intake.32

Several issues pertaining to the study design warrant comment. Strengths include the naturalistic setting of a food court for implementing the feeding protocols and a crossover design for hypothesis testing, thereby minimizing the possibility of confounding by demographic and behavioral variables. Limitations include a small sample size, restricting generalizability, and evaluation of only 1 combination of fast food menu items.

CONCLUSIONS

Portioning menu items and slowing the eating rate did not decrease energy intake during a fast food meal. Our findings suggest that excess energy intake during an extra-large fast food meal is not attributable simply to distorted visual cues regarding consumption of foods and beverage or inadequate time for development of satiety signals. Rather, inherent characteristics related to the nutritional composition of conventional fast food likely promote excess energy intake. Although we cannot dismiss the potential public health benefits of strategies aimed at portion control, such as eliminating extra-large meals from fast food menus, fundamental improvements in the nutritional quality of fast food may be necessary to diminish the reported contribution of fast food to energy intake and risk for obesity.2,25,37

ACKNOWLEDGMENTS

This study was supported by grant P30 DK40561 awarded by the National Institutes of Health (Bethesda, MD) to support the Harvard Clinical Nutrition Research Center; the Charles H. Hood Foundation (Boston, MA); grant M01 RR02172 awarded by the National Institutes of Health (Bethesda, MD) to support the General Clinical Research Center at Children’s Hospital Boston (Boston,
REFERENCES


9. Levitsky DA, Youn T. The more food young adults are served, the more they overeat. *J Nutr.* 2004;134:2546–2549


26. Brownell KD, Horgen KB. *Food Fight: The Inside Story of the Food Industry, America’s Obesity Crisis, and What We Can Do About It.* Chicago, IL: Contemporary Books; 2004


A Cluster-Randomized Trial of Benchmarking and Multimodal Quality Improvement to Improve Rates of Survival Free of Bronchopulmonary Dysplasia for Infants With Birth Weights of Less Than 1250 Grams

Michele Walsh, MD, MS*, Abbott Laptook, MD*, S. Nadya Kazzi, MD, MPH*, William A. Engle, MD*, Qing Yao, PhD*, Maynard Rasmussen, MD*, Susie Buchter, MD*, Gregory Heldt, MD*, William Rhine, MD*, Rose Higgins, MD*, Kenneth Poole, PhD*, for the National Institute of Child Health and Human Development Neonatal Research Network

*Department of Pediatrics, Rainbow Babies & Children's Hospital, Case Western Reserve University, Cleveland, Ohio; *Department of Pediatrics, Women & Infants Hospital, Brown University, Providence, Rhode Island; *Department of Pediatrics, Wayne State University, Detroit, Michigan; *Department of Pediatrics, Riley Hospital for Children, Indiana University, Indianapolis, Indiana; *Research Triangle Institute, Research Triangle Park, North Carolina; *Department of Pediatrics, Sharp Mary Birch Hospital, University of California, San Diego, San Diego, California; *Department of Pediatrics, Emory University, Atlanta, Georgia; *Department of Pediatrics, University of California, San Diego, San Diego, California; *Department of Pediatrics, Lucille Packard Children's Hospital, Stanford University, Palo Alto, California; *National Institute of Child Health and Human Development, Rockville, Maryland

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

ABSTRACT

OBJECTIVE. We tested whether NICU teams trained in benchmarking and quality improvement would change practices and improve rates of survival without bronchopulmonary dysplasia in inborn neonates with birth weights of <1250 g.

METHODS. A cluster-randomized trial enrolled 4093 inborn neonates with birth weights of <1250 g at 17 centers of the National Institute of Child Health and Human Development Neonatal Research Network. Three centers were selected as best performers, and the remaining 14 centers were randomized to intervention or control. Changes in rates of survival free of bronchopulmonary dysplasia were compared between study year 1 and year 3.

RESULTS. Intervention centers implemented potentially better practices successfully; changes included reduced oxygen saturation targets and reduced exposure to mechanical ventilation. Five of 7 intervention centers and 2 of 7 control centers implemented use of high-saturation alarms to reduce oxygen exposure. Lower oxygen saturation targets reduced oxygen levels in the first week of life. Despite these changes, rates of survival free of bronchopulmonary dysplasia were all similar between intervention and control groups and remained significantly less than the rate achieved in the best-performing centers (73.3%).

CONCLUSIONS. In this cluster-randomized trial, benchmarking and multimodal quality improvement changed practices but did not reduce bronchopulmonary dysplasia rates.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2656
doi:10.1542/peds.2006-2656

Key Words
quality improvement, randomized trial, bronchopulmonary dysplasia

Abbreviations
QI—quality improvement
BPD—bronchopulmonary dysplasia
PMA—postmenstrual age
CPAP—continuous positive airway pressure
OR—odds ratio
CI—confidence interval

Accepted for publication Dec 27, 2006
Address correspondence to Michele Walsh, MD, MS, Rainbow Babies & Children's Hospital, 11100 Euclid Ave, Mailstop 6010, Cleveland, OH 44106-6010. E-mail: michele.walsh@cwru.edu

PECIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
In 2000, publication of the seminal report To Err Is Human by the Institute of Medicine rocked the medical world and shook the confidence of patients in the health care system. The Institute of Medicine then examined deficiencies in common health care practices in the report Crossing the Quality Chasm and found that health care processes were subject to significant deviations from expected practice that led to suboptimal care and poor patient outcomes. Quality improvement (QI) techniques, adapted from industry, have been used to improve care and patient outcomes. However, the National Academies of Sciences found that research on QI lagged behind real-world applications. The Institute of Medicine concluded that most health care professionals were trained inadequately in QI. The optimal techniques for training teams and changing practice are not known.

One technique in QI is benchmarking, in which best-performing centers are identified and practices are examined and emulated at other centers to improve outcomes. The rationale is that institutions with excellent performance for a given outcome apply specific clinical practices that are most effective. They may also display structural or cultural organizational features that contribute to excellent outcomes. By visiting these centers and reviewing the evidence in the literature, teams from other institutions can identify these practices and organizational features. Then, by applying methods learned in QI training, the teams should be able to implement the identified practices and to modify their organizations in ways that lead to better outcomes. Although such QI teams are used increasingly in health care organizations, their efficacy has not been evaluated rigorously. Evaluations of the impact of QI teams have revealed mixed results. Shojania et al reviewed systematically the results of QI studies and found evidence for only modest effects, together with evidence of publication bias, with small trials being more likely to demonstrate positive results and larger trials more often yielding negative results. Also clouding the evaluation of QI is the fact that teams are often self-selected and highly motivated. It is not known whether the successful improvements in outcomes reported by such teams can be generalized when applied on a wider scale. In addition, it is unclear whether teams can select from the potentially thousands of clinical practices used to identify a subset of practices that can be applied to produce improved outcomes.

Survival rates for very low birth weight neonates (<1500 g) have improved steadily, with 83% of all such neonates surviving. Although most of the survivors are healthy, many develop a chronic lung injury, bronchopulmonary dysplasia (BPD), which is a significant health burden.

In 1998, 55% of neonates with birth weights of <1250 g who were born at centers in the National Institute of Child Health and Human Development Neonatal Research Network either died or developed BPD. The incidence rates of BPD vary by center and are not explained by differences in birth weight, gestational age, race, frequency of prenatal steroid use, or incidence of respiratory distress syndrome. Therefore, differences in treatment practices may contribute to the development of BPD.

We conducted a cluster-randomized, controlled trial to test whether NICUs trained in benchmarking and multimodal QI techniques could improve rates of survival without BPD for neonates with birth weights of <1250 g, compared with centers with usual practice.

**METHODS**

**Participants**

Seventeen centers of the National Institute of Child Health and Human Development Neonatal Research Network participated in the trial, with practices analyzed for inborn neonates with birth weights of <1250 g. In January 2001, the 3 centers with the highest rates of survival free of BPD (top 3 performers in 1998–2000; rate of survival free of BPD: 62.5%) were identified as the benchmark centers (see “Acknowledgments”).

**Randomization**

Our intention was to improve the use of potentially better practices by the entire neonatal care team. The NICU, rather than the patient, was the unit of randomization, because the intervention was applied to a team representing the NICU. In June 2001, the 14 remaining centers were assigned randomly, with computer-generated codes prepared in sealed opaque envelopes by the data center, to the intervention group (N = 7) or the control group (N = 7). Envelopes were distributed in person, and all were opened simultaneously. A flow diagram of study participants and units is shown in Fig 1.

**Interventions**

**Data Collection**

Before randomization, all 14 eligible sites selected multidisciplinary teams (neonatologist, neonatal nurse, and respiratory therapist); members were respected clinical experts at their sites. From January to June 2001, data on preintervention practices were collected at the benchmark centers and at each intervention center and were analyzed by each intervention team to identify care differences. From June to November 2001, teams conducted self-study and literature review by using the preintervention data.

**Training in QI Practices**

All team members attended an 8-hour training session on QI led by a team of experts (see “Acknowledgments”). Sessions introduced systems thinking, cycles of rapid change, measurement tools, and the concept of potentially better practices. Teams were provided with literature reviews of care practices, including published
meta-analyses and reviews from the Cochrane Collaboration. Core teams met face to face on 2 occasions and then in teleconferences. Teleconferences at 4- to 8-week intervals throughout the 2-year intervention period supported initial training. One control site and 2 intervention sites had participated in previous rapid-cycle QI processes.

Site Visits and Selection of Interventions
The teams from the intervention centers visited each benchmark center in November 2001. Benchmark centers delivered a presentation on their self-assessments of practices responsible for their high rates of survival free of BPD to the intervention teams. Teams also observed care directly at each benchmark site, collaborated in document care practices, and compared these with the benchmark self-assessments. In addition, intervention teams scrutinized extensive data collected by research nurses during the preintervention period at the benchmark centers and at their own centers.

From these data-driven assessments, teams identified 27 potentially better practices at the benchmark centers, in 3 domains, namely, delivery room care, ventilation practices, and nutrition and fluid practices (Table 1).

Two other domains evaluated originally, that is, infection rates and infection control practices and organizational structure, were not different at the better-performing centers and were not selected for implementation. Overall care in the benchmark centers was characterized by ventilation with lower tidal volumes (2 centers with emphasis on use of nasal continuous positive airway pressure [CPAP] and 1 center with continued mechanical ventilation with low tidal volumes) and lower oxygen saturation targets.

After the site visits, the teams reviewed published evidence, focusing on systematic reviews, evaluated the quality of the evidence by using the criteria of the Oxford Centre for Evidence-Based Medicine, and collaborated with colleagues at their centers to identify practices at the benchmark centers that were different from those in their own centers. The core team members led conferences to develop consensus with their colleagues and together selected potentially better practices for implementation. Because preexisting practices differed at each center (by design), each unit developed a unique set of interventions based on their local practice patterns; however, many centers chose similar interventions. Intervention centers chose between 5 and 13 potentially better practices per center (median: 7 practices) for im-

---

FIGURE 1
Flow diagram of participating centers and neonates in the trial. GA indicates gestational age.
Implementation. Specific practices at the benchmark centers and those selected by the intervention centers are shown in Table 1. More-detailed information on the potentially better practices, levels of evidence, and metrics used is contained in the Appendix.

Implementation of all selected potentially better practices was tracked with statistical control charts. Each intervention was assigned a predefined method of objective measurement based on observation of practices at the better-performing centers. For example, implementation of the use of high-saturation alarms was tracked with random audits of the use of alarms at the intervention sites and the control sites. The metrics used to track practice changes are summarized in the Appendix. Control charts (example shown in Fig 2) were generated by the data center with SAS software (SAS Institute, Cary, NC), provided to the intervention teams at 4- to 6-week intervals throughout the 2-year intervention period, and shared with all members of the NICUs to reinforce practice changes. Each team received control charts of all practices selected for implementation by any team. In this study, successful implementation of an intervention

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Summary of Potentially Better Practices and Selections According to Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>Benchmark A</td>
</tr>
<tr>
<td>Delivery room</td>
<td></td>
</tr>
<tr>
<td>1. Fellow or attending physician at delivery</td>
<td>P</td>
</tr>
<tr>
<td>2. Respiratory therapist at delivery</td>
<td>P</td>
</tr>
<tr>
<td>3. Consistent equipment in all delivery rooms</td>
<td>P</td>
</tr>
<tr>
<td>5. Prophylactic use of surfactant</td>
<td>P</td>
</tr>
<tr>
<td>6. Use of device to provide positive end-expiratory pressure</td>
<td>P</td>
</tr>
<tr>
<td>Respiratory care</td>
<td></td>
</tr>
<tr>
<td>7. Selective intubation with liberal use of CPAP</td>
<td></td>
</tr>
<tr>
<td>8. Early use of surfactant if intubated</td>
<td>P</td>
</tr>
<tr>
<td>9. Assessment of volume/pressure and targeting of lowest levels to achieve modest chest rise if intubated</td>
<td>P</td>
</tr>
<tr>
<td>10. Aggressive weaning and early extubation if intubated</td>
<td>P</td>
</tr>
<tr>
<td>11. Higher PaCO₂ targets for all patients</td>
<td>P</td>
</tr>
<tr>
<td>12. Lower oxygen saturation goals</td>
<td>P</td>
</tr>
<tr>
<td>13. High-saturation alarm set at 95%</td>
<td>P</td>
</tr>
<tr>
<td>15. Avoidance of hand-bagging for patients undergoing ventilation</td>
<td>P</td>
</tr>
<tr>
<td>17. Prophylactic use of methylxanthines before extubation</td>
<td>P</td>
</tr>
<tr>
<td>18. Consensus regarding ventilatory management</td>
<td>P</td>
</tr>
<tr>
<td>Nutrition/fluids</td>
<td></td>
</tr>
<tr>
<td>19. Limited intravenous fluids</td>
<td>P</td>
</tr>
<tr>
<td>20. High-humidity environments</td>
<td>P</td>
</tr>
<tr>
<td>21. Limited volume expansion to treat low blood pressure</td>
<td>P</td>
</tr>
<tr>
<td>22. Aggressive approach to patent ductus arteriosus</td>
<td>P</td>
</tr>
<tr>
<td>25. Full total parenteral nutrition with increasing enteral feeding</td>
<td>P</td>
</tr>
<tr>
<td>26. Frequent use of human milk</td>
<td>P</td>
</tr>
<tr>
<td>27. Vitamin A prophylaxis</td>
<td>P</td>
</tr>
</tbody>
</table>

P indicates a center already practicing this standard; X, an intervention selected by the center.
was defined as a statistically significant improvement from preintervention performance achieved within the 2-year intervention period and maintained at the end of the period. Data on practices in the best-performing centers before the study were available to the teams, but data on performance in the control centers were masked. Control centers were provided with annual summary data provided routinely to all network centers and were masked with respect to the work at the intervention centers. Control centers were prohibited from participating in other QI collaborative efforts focused on BPD. Because of ethical concerns with prohibiting practice changes in control centers during the 3-year trial period, local quality efforts initiated by clinicians who were not members of the research team were permitted. One control center changed its management approach to deemphasize mechanical ventilation and to emphasize nasal CPAP during the trial.

**Outcomes**

The primary outcome measure was the change in survival free of BPD between year 1 and year 3. BPD was assessed at postmenstrual age (PMA) of 36 weeks by using a validated physiologic definition that combined respiratory support and oxygen saturation and was developed for this trial. Infants who were discharged before 36 weeks were assigned the diagnosis of BPD if discharged from the hospital with oxygen. For infants who were transferred before 36 weeks, a hierarchy was used to determine the diagnosis of BPD. If possible, a room air challenge was performed and the infant was assigned the diagnosis of BPD on the basis of the results of the challenge. If a challenge was not possible, then the receiving institution was contacted and the infant was assigned the outcome of BPD if he or she was receiving oxygen supplementation, CPAP, or ventilation. If no information about the status at 36 weeks was available, then the infant was assigned the diagnosis of BPD if he or she was receiving oxygen supplementation, CPAP, or ventilation at the time of transfer. Overall, the outcome of BPD was determined for 99.9% of infants at intervention sites and 98.8% of infants at control sites. Secondary outcomes included death before hospital discharge, BPD severity (assessed with a modification of the National Institutes of Health consensus definition of BPD that included the physiologic definition), durations of mechanical ventilation, CPAP, and oxygen use, and length of hospital stay. Other measures of common neonatal comorbidities were specified before the trial began and included severe intraventricular hemorrhage (Papile stage III or IV), cystic periventricular leukomalacia, severe retinopathy of prematurity (stage 3 or more), pneumothorax, patent ductus arteriosus, necrotizing enterocolitis (stage 2 or more), and late-onset sepsis (positive blood culture at >72 hours of age). Arterial oxygen values, together with complete blood gas data and respiratory support information, were measured every 6 hours on days 1 to 7, at the values closest to 6 AM, noon, 6 PM, and midnight. The information was also recorded on days 14, 21, and 28 of life. All values were averaged. Severity of illness was assessed at 24 hours of age by using the Score of Neonatal Acute Physiology II. Neonatal research nurses abstracted all data by using standardized definitions. Data were entered remotely.
through electronic submission. Quality control procedures included range checking, internal comparisons for logic violations, and comparison of expected and observed values.

**Human Subject Protection**

The institutional review board at every site approved the study. One center provided families with a letter of information, and all others were given a waiver of consent requirements to collect deidentified data. The trial was registered at inception with the US National Library of Medicine trial registry (trial registration NCT00067613 [see www.clinicaltrials.gov]). In April 2003, data were reviewed by an independent data monitoring and safety committee, which recommended trial continuation.

**Study Time Line and Statistical Methods**

The preintervention year (study year 1) began in March 2001. Centers were assigned randomly to intervention or control in June 2001, to permit centers to free investigators for site visits in October and November 2001. Centers selected interventions, began implementation in May 2002, and continued interventions through a 2-year intervention period (study year 2 and year 3). Outcomes were compared between year 1 (March 2001 to May 2002) and year 3 (April 2003 to May 2004).

Analyses included all neonates with birth weights of <1250 g who were born at the centers and were free of major malformations. All analyses were based on an intent-to-treat model according to center assignment. Comparisons of the intervention and control centers were assessed by using mixed-model methods (SAS 9.1 loadable PROC GLIMMIX for binary outcomes and PROC MIXED for continuous outcomes), with the center entered as a random effect. These analyses accounted for the intraclass correlation within each center attributable to clustering from randomization according to center. The model for the analysis included the following terms: group (intervention or control), study year (year 1 or year 3), and group-year interaction. The group-year interaction term measures the difference between the 2 groups in changes from year 1 to year 3, which is the parameter of interest. A 0 coefficient for this term indicates a null treatment effect between the 2 groups. Other characteristics of the infants present at the time of birth were added to the model, including birth weight, gestational age of <26 weeks, race, gender, and prenatal steroid exposure. An interaction between gestational age (<26 weeks versus ≥26 weeks) and the main effect (group-year interaction) was added to the model if it showed a significant effect. Summary outcomes are shown when the gestational age interaction was not significant, and outcomes according to gestational age group are shown when the interaction was significant. Binary outcomes are presented as adjusted odds ratios (ORs) for year 3 versus year 1, with 95% confidence intervals (CIs), and continuous outcomes as the adjusted difference of year 3 versus year 1. A model with a term for severity of illness that included the Score for Neonatal Acute Physiology II with perinatal extension was also studied; the results obtained by using the additional term for severity of illness were not different from the results of the first model and are not shown. We prespecified secondary analyses that evaluated the impact of the intervention according to center and according to gestational age (26 weeks versus ≥26 weeks). All study analyses were completed with SAS 9.1 (SAS Institute).

**Sample Size**

In this trial, with the center rather than the individual patient as the unit of randomization, sample size calculations accounted for both interhospital and intrahospital variability. The comparison of interest was the intervention/control difference in the change in rates of survival free of BPD between study years 1 and 3. The methods of Gail et al from the Community Intervention Trial for Smoking Cessation were used in these calculations. Based on network data for 1999 and 2000, the rate of survival free of BPD (defined as oxygen use at 36 weeks) in inborn neonates of <1250 g was 45%. We calculated a sample size of 1400 neonates in each year of the study (100 patients per center in 14 centers) to yield 80% power (2-sided α = .05) to detect an absolute difference of 14% in the change in rates of survival free of BPD between year 1 and year 3 in the intervention versus control groups. The magnitude of the effect selected was relatively large and was based on effect size seen in the only published work of collaborative QI to reduce BPD.

**RESULTS**

**Patient Population and Care Practices Before Intervention**

The population for these analyses included 4095 live-born neonates with birth weights of 401 to 1250 g who were born at the 14 randomized centers between March 1, 2001, and April 30, 2004. Neonates who were not born at the centers and those with major malformations were excluded. Two neonates with major malformations were enrolled incorrectly and were removed subsequently, leaving 4093 neonates in the cohort. Among the 4093 neonates, 2871 were from year 1 and year 3 of the study and therefore were included in the analysis. The hospitals in the study arms were all level IIIB units with large volumes and accredited residency and neonatal/perinatal programs. Centers randomized to the control group were larger than those randomized to the intervention group, leading to more infants in the control centers (Fig 1). Infants in the intervention centers were slightly larger, were more mature, and included fewer white infants, compared with those in the control
centers (Table 2). Infants in the intervention centers were born to mothers who had received less prenatal steroid treatment (78% vs 87%; P < .0001), and the infants had higher severity of illness scores, as measured with the Score of Neonatal Acute Physiology II (23.8 ± 15.6 vs 20.9 ± 15.4; P = .0002). Despite these differences in infant characteristics, the incidence rates of BPD at 36 weeks, measured with the physiologic definition, were similar for the intervention and control centers (25.7% vs 28.3%; P = .30), as were the incidence rates of oxygen use at 36 weeks (38.5% vs 36.1%; P = .40).

Characteristics of the infants within centers did not differ across the 3 years of the trial with respect to birth weight, gestational age, gestational age of <24 weeks, gender, or prenatal steroid exposure. The only attribute with a statistically significant change was an increase in the percentage of black infants born at centers randomized to the intervention group from 30.6% to 36.3% (P = .047).

Interventions

The majority of intervention centers implemented their selected practices, although the rate of success did vary according to center (median rate of success: 75%; range: 40%–100%). The intervention group did change respiratory care practices (Table 3). Both intervention and control centers decreased the duration of delivery of the first surfactant dose, with intervention centers decreasing from a median of 51 minutes to 31 minutes and control centers decreasing from a median of 41 minutes to 21 minutes. The intervention group significantly increased the use of CPAP on the first day of life in the intervention centers, from 4.0 /H11005 to 2.7 days in year 1 and 117 mL/kg per day in year 3; in contrast, the values for the control centers were similar in years 1 and 3 (64.3 ± 26.4 mm Hg). Rates of postnatal steroid use declined significantly in both groups (intervention: from 14.3% to 4.4%; P < .01; control: from 14.3% to 5.6%; P < .01).

The final practice changes selected by intervention centers were restrictions of intravenous fluid volumes. Four centers selected this intervention, and 3 were successful in implementation, with reductions in delivered intravenous fluid (intervention: 126.8 mL/kg per day on day 3 of life in year 1 and 117 mL/kg per day in year 3; control: 125.1–122.6 mL/kg per day).

Outcomes

Intervention centers did not change the frequency of survival free of BPD faster than control centers (change score for intervention group adjusted for clustering: change significantly (3.5 ± 2.8 days versus 3.4 ± 2.8 days). Despite having persistently higher rates of intubation on day 1 of life than did control centers between study year 1 and study year 3, intervention centers decreased the total duration of respiratory support by 5.3 days, whereas the control group decreased the duration by 4.1 days.

Intervention centers also implemented policies to reduce target oxygen saturations more frequently than did control centers. Five (71%) of 7 intervention centers and 2 (28.5%) of 7 control centers implemented use of high-saturation alarms to reduce oxygen saturation exposure. In monthly audits at all centers, patients receiving oxygen at intervention centers were more likely to have a high-saturation alarm in use (68.7% of audits with alarm in use; range: 39.8%–100%) than were those at control centers (10.5%; range: 0%–33%). This contributed to a reduction in the arterial oxygen levels measured in the first week of life in the intervention centers, from 74.1 ± 33.6 mm Hg (mean ± SD) in year 1 to 62.7 ± 21.2 mm Hg in year 3. In contrast, the values for the control centers were similar in years 1 and 3 (62.2 ± 30 vs 64.3 ± 26.4 mm Hg). Rates of postnatal steroid use declined significantly in both groups (intervention: from 14.3% to 4.4%; P < .01; control: from 14.3% to 5.6%; P < .01).

Comparisons of population characteristics in each group between study year 1 and study year 3 were not significantly different, except for race in the intervention centers.

| TABLE 2 | Description of Populations at Intervention and Control Centers |
|------------------|------------------|------------------|------------------|------------------|
| Characteristic   | Intervention Centers (N = 1752) | Control Centers (N = 2341) |
|                  | Year 1 (N = 625) | Year 3 (N = 595) | Year 1 (N = 787) | Year 3 (N = 864) |
| Birth weight, mean ± SD, g | 932 ± 207 | 922 ± 206 | 899 ± 209 | 912 ± 212 |
| PMA, mean ± SD, wk | 273 ± 2.3 | 273 ± 2.5 | 270 ± 2.3 | 271 ± 2.5 |
| PMA of <26 wk, n (%) | 151 (24.2) | 157 (26.4) | 229 (28.8) | 244 (28.2) |
| Male gender, n (%) | 338 (54.1) | 311 (52.3) | 394 (49.9) | 432 (50.2) |
| Race, n (%) | | | | |
| White | 196 (31.4)* | 170 (28.6)* | 329 (41.8) | 365 (42.3) |
| Black | 191 (30.6) | 216 (36.3) | 368 (46.8) | 377 (43.6) |
| Other | 238 (38.1) | 209 (35.1) | 90 (11.4) | 122 (14.2) |
| Maternal age, mean ± SD, y | 27.2 ± 6.4 | 27.3 ± 6.7 | 27.2 ± 6.4 | 27.0 ± 6.3 |
| Prenatal steroid exposure, n (%) | 482 (77.1) | 470 (79.1) | 680 (86.4) | 761 (88.2) |
| Score of Neonatal Acute Physiology II, mean ± SD | 23.4 ± 15.0 | 24.6 ± 16.8 | 22.4 ± 14.6 | 19.5 ± 15.7 |

Comparisons of population characteristics in each group between study year 1 and study year 3 were not significantly different, except for race in the intervention centers.

* P = .047.
Effects of Intervention According to Gestational Age

During study design, we hypothesized that the intervention would be less successful for less-mature infants because of increased biological vulnerability. Therefore, we analyzed outcomes for 2 gestational age groups, that is, PMA of <26 weeks and PMA of ≥26 weeks. As anticipated, for infants with PMA of ≥26 weeks across the 3 years of the trial, rates of survival to PMA of 36 weeks (94% vs 66%) and survival free of BPD (74% vs 30%) were significantly higher than those for infants with PMA of <26 weeks. Overall, results did differ according to PMA group (Fig 4). For infants with PMA of ≥26 weeks, there were no differences in rates of survival to PMA of 36 weeks between intervention and control centers, but a statistically nonsignificant trend toward increased rates of severe intraventricular hemorrhage in the intervention centers between year 1 and year 3 (14.4% vs 18.1%), compared with the control centers (13.3% vs 14.1%). The incidence rates of periventricular leukomalacia, severe retinopathy of prematurity, necrotizing enterocolitis, patent ductus arteriosus, and growth failure were similar between the 2 groups in study years 1 and 3.

Outcomes According to Center

Results did differ according to center in both the intervention and control groups. In the intervention group, 1 center had a significantly improved rate of survival free of BPD (OR: 1.93; 95% CI: 1.00–3.74) and 6 centers had no significant change (OR: 0.37–1.22). However, in 2 of the centers with no significant change overall, there was a significant interaction between the intervention and PMA. In those 2 centers, no change in BPD rates was seen for neonates born at PMA of ≥26 weeks (OR: 1.33; 95% CI: 0.75–2.36) but worsened outcomes were seen for infants born at PMA of <26 weeks (OR: 0.18; 95% CI: 0.06–0.55). Those 2 centers drove the gestational age differences described in the overall trial. In the control centers, 1 center significantly improved the rate of survival free of BPD (OR: 1.96; 95% CI: 1.00–3.84), whereas 5 centers showed no significant change (OR: 0.40–1.35) and 1 center worsened significantly (OR: 0.53; 95% CI: 0.29–0.99).

DISCUSSION

We showed that, in a rigorous trial using the center as the unit of randomization, centers that implemented QI processes were successful in changing care practices but did not improve rates of survival free of BPD for high-

---

**TABLE 3** Respiratory Outcomes for Intervention and Control Groups

<table>
<thead>
<tr>
<th>Variables</th>
<th>Benchmark Centers, Year 1</th>
<th>Intervention Centers</th>
<th>Control Centers</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year 1</td>
<td>Year 3</td>
<td>Change Score/ OR (95% CI), Year 3 vs Year 1*</td>
<td>Year 1</td>
</tr>
<tr>
<td>Continuous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total duration of ventilation, mean ± SD, d</td>
<td>13.8 (1.97)</td>
<td>15.0 (1.82)</td>
<td>14.3 (1.93)</td>
<td>-1.2 (-3.1 to 0.7)</td>
</tr>
<tr>
<td>Duration of ventilation in first 7 d, mean ± SD, d</td>
<td>3.5 (2.8)</td>
<td>4.0 (2.7)</td>
<td>3.5 (2.8)</td>
<td>-0.6 (-0.8 to -0.3)</td>
</tr>
<tr>
<td>Duration of CPAP, mean ± SD, d</td>
<td>6.6 (9.5)</td>
<td>8.5 (11.9)</td>
<td>8.0 (11.4)</td>
<td>9.4 (11.2)</td>
</tr>
<tr>
<td>Duration of oxygen therapy, mean ± SD, d</td>
<td>28.1 (26.7)</td>
<td>33.2 (27.0)</td>
<td>29.0 (27.0)</td>
<td>-4.9 (-7.6 to -2.2)</td>
</tr>
<tr>
<td>Dichotomous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPAP use on day 1 of life, %</td>
<td>15.8</td>
<td>16.9</td>
<td>24.2</td>
<td>1.7 (1.27–2.30)</td>
</tr>
<tr>
<td>Surfactant use, %</td>
<td>66.6</td>
<td>75.8</td>
<td>76.9</td>
<td>0.96 (0.71–1.30)</td>
</tr>
<tr>
<td>Pneumothorax, %</td>
<td>6.7</td>
<td>8.1</td>
<td>7.4</td>
<td>0.85 (0.55–1.31)</td>
</tr>
</tbody>
</table>

See “Methods” for elements included in the model

* For continuous variables, the model compared the differences in the change scores between year 1 and year 3 for intervention and control groups; for dichotomous variables, the model compared the differences in the ORs.
risk preterm neonates, compared with control centers. Intervention centers benefited from working together over time, tracking their results, and changing practices. Intervention centers modified their practices to be more similar to those of the benchmark centers. Intervention centers significantly decreased measured oxygen tension in the first week of life and significantly increased the use of nasal CPAP, compared with control centers. Despite these changes in delivery room and respiratory care practices, we did not demonstrate increased rates of survival free of BPD. In part, unexpectedly low rates of BPD in the preintervention year in the intervention centers made improvements difficult.

By chance, centers randomized to intervention had higher rates of endotracheal intubation (77% vs 66%), lower rates of CPAP on day 1 of life (16.9% vs 26.5%), and more days of mechanical ventilation in the first 7 days of life (4.0 ± 2.7 vs 3.5 ± 2.8 days). The goal of the study was to assess the utility of benchmarking to accelerate practice changes. As is normal in QI studies, intervention centers were focused initially on the performance of the 3 benchmark centers and later on performance at their own centers. Intervention centers were masked with respect to processes at the control centers. Differences in practices at centers were anticipated in the design and represent the reason why a change score between year 1 and year 3 was used as the primary outcome statistic; such an approach corrects for potential baseline differences in patient characteristics and practices. Despite differences in practices, interven-

TABLE 4  Primary Outcomes for Benchmark, Intervention, and Control Groups

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Benchmark Centers, Year 1</th>
<th>Intervention Centers</th>
<th>Control Centers</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survival free of BPD*</td>
<td>73.3</td>
<td>63.3</td>
<td>62.2</td>
<td></td>
</tr>
<tr>
<td>Gestational age of &lt;26 wk</td>
<td>52.8</td>
<td>38.4</td>
<td>22.3</td>
<td>0.47 (0.27–0.82)</td>
</tr>
<tr>
<td>Gestational age of ≥26 wk</td>
<td>83.1</td>
<td>71.3</td>
<td>76.5</td>
<td>1.23 (0.88–1.71)</td>
</tr>
<tr>
<td>Survival to PMA of 36 wk</td>
<td>84.4</td>
<td>85.3</td>
<td>84.2</td>
<td>0.96 (0.67–1.37)</td>
</tr>
<tr>
<td>BPD, physiologic definition</td>
<td>13.4</td>
<td>25.7</td>
<td>26.1</td>
<td></td>
</tr>
<tr>
<td>Gestational age of &lt;26 wk</td>
<td>24.0</td>
<td>37.6</td>
<td>60.7</td>
<td>2.53 (1.31–4.90)</td>
</tr>
<tr>
<td>Gestational age of ≥26 wk</td>
<td>9.5</td>
<td>23.1</td>
<td>18.7</td>
<td>0.80 (0.56–1.16)</td>
</tr>
<tr>
<td>BPD severity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>9.8</td>
<td>11.3</td>
<td>12.2</td>
<td>14.8 (13.0)</td>
</tr>
<tr>
<td>Moderate</td>
<td>3.3</td>
<td>14.3</td>
<td>13.8</td>
<td>12.8 (13.9)</td>
</tr>
<tr>
<td>Mild</td>
<td>0.8</td>
<td>8.1</td>
<td>4.4</td>
<td>4.5 (4.7)</td>
</tr>
<tr>
<td>None</td>
<td>86.2</td>
<td>66.2</td>
<td>69.5</td>
<td>67.0 (67.9)</td>
</tr>
<tr>
<td>Undefined</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.8 (0.4)</td>
</tr>
</tbody>
</table>

Benchmark center data were collected from October 2000 to May 2001. All analyses compared the differences in the changes between study year 1 and year 3 in the intervention centers, compared with the control centers. Variables included in the model were group assignment (intervention versus control), study year, group-study year, birthweight, gender, race, prenatal steroid exposure, interaction with gestational age of <26 weeks, and random clustering effect according to center.

* BPD severity was defined with a modification of the National Institutes of Health consensus definition of BPD that included results of a timed room air challenge (see “Methods”). For this analysis, BPD severity was dichotomized as any BPD versus no BPD.

** Analyses that demonstrated effect modification by gestational age are shown with overall results and with results according to gestational age group. The significance of interaction term is shown in parentheses.
tion centers did not improve rates of survival free of BPD faster than control centers. If in fact the practices on which the centers focused were important determinants of BPD, then the intervention centers should have been biased toward a greater likelihood of finding an improvement, because they had greater opportunities to improve practice.

BPD is a complex disease with multifactorial pathophysiologic processes. Disruption of the development of fragile lung parenchyma at a critical period of alveolar and vasculature maturation is thought to be a primary determinant of BPD. Key contributors to such disruption are oxidative injury from oxygen, ventilator-induced lung injury, and inflammation. Interventions in this trial reduced ventilator pressures, reduced time with mechanical ventilation, and reduced oxygen concentrations. Interventions to avoid volutrauma were supported by evidence from preclinical studies in animals and human neonates and were reinforced by data and direct observation at the benchmark centers. However, as shown in the Appendix, the strength of the evidence for these interventions in human neonates is weak, with few randomized trials. Existing evidence equally supports 2 philosophically divergent interventions, that is, endotracheal intubation and the early delivery of surfactant versus avoidance of endotracheal intubation and the use of CPAP. It is currently unclear which is the superior approach. Concern regarding potential injury from oxidative stress from high-oxygen environments for which preterm neonates are poorly prepared biologically led centers to emphasize oxygen reduction, the only intervention selected by all 7 intervention sites. In a posthoc analysis using both intervention and control centers, interventions focused on pressure reduction rather than oxygen reduction were more successful in reducing BPD rates (data not shown).

Although interest in QI to improve health care and outcomes is not new, rigorous randomized trials evaluating the method have been conducted only recently, with mixed results. The majority of the trials focused on improving delivery of evidence-based services to adults (eg, use of β-receptor blockers) or improving the efficiency of care delivery (eg, reducing waiting times). Several trials tested the utility of multimodal interventions. In a cluster-randomized trial, Ferguson et al demonstrated improved preoperative β-receptor blocker and internal mammary graft use in a nationwide QI effort. Although a statistically significant increase in prescription of preoperative β-receptor blocker therapy was seen, the magnitude of the increase was modest (7.3% vs 3.6% in control centers). Mehta et al tested a multimodal intervention led by local opinion leaders to measure the impact on 11 indicators of the quality of acute myocardial infarction care. Some indicators improved in intervention centers, whereas others improved more in the control centers. Overall, the absolute gains ranged

### TABLE 5 Other Neonatal Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Intervention Centers</th>
<th>Control Centers</th>
<th>Adjusted P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% OR (95% CI), Year 3 vs Year 1</td>
<td>% OR (95% CI), Year 3 vs Year 1</td>
<td></td>
</tr>
<tr>
<td>Severe intraventricular hemorrhage</td>
<td>14.4 18.1</td>
<td>13.3 14.1</td>
<td>0.28 (0.03) a</td>
</tr>
<tr>
<td>PMA of &lt;26 wk</td>
<td>26.2 45.6</td>
<td>26.0 28.6</td>
<td>1.12 (0.72–1.75)</td>
</tr>
<tr>
<td>PMA of ≥26 wk</td>
<td>10.8 8.8</td>
<td>8.4 8.7</td>
<td>1.05 (0.68–1.60)</td>
</tr>
<tr>
<td>Periventricular leukomalacia</td>
<td>1.5 1.9</td>
<td>2.2 1.5</td>
<td>0.68 (0.32–1.47)</td>
</tr>
<tr>
<td>Retinopathy of prematurity, stage ≥3</td>
<td>13.3 16.2</td>
<td>16.3 17.0</td>
<td>1.10 (0.79–1.54)</td>
</tr>
<tr>
<td>Necrotizing enterocolitis, stage ≥2</td>
<td>7.8 8.9</td>
<td>9.1 10.3</td>
<td>1.19 (0.85–1.66)</td>
</tr>
<tr>
<td>One or more of above morbidities</td>
<td>84.6 84.3</td>
<td>84.8 83.7</td>
<td>0.97 (0.72–1.32)</td>
</tr>
<tr>
<td>Discharged with oxygen</td>
<td>19.6 12.6</td>
<td>8.4 9.2</td>
<td>34.02)</td>
</tr>
<tr>
<td>Gestational age of &lt;26 wk</td>
<td>25.8 28.6</td>
<td>24.8 30.2</td>
<td>1.28 (0.65–2.52)</td>
</tr>
<tr>
<td>Gestational age of ≥26 wk</td>
<td>15.6 10.6</td>
<td>5.1 5.1</td>
<td>0.94 (0.52–1.68)</td>
</tr>
</tbody>
</table>

For dichotomous variables, the model compared the differences in the ORs for year 3 versus year 1 for the intervention and control groups. See “Methods” for elements included in the model. a Variable shows significant effect modification by gestational age (P < .03); overall results and results according to gestational age group are shown. The significance of interaction term is shown in parentheses.

![Figure 4](https://example.com/figure4.png)

**FIGURE 4**
Impact of PMA on rates of BPD in study years 1 and 3 in the intervention and control centers. Intervention centers demonstrated a low rate of BPD in the preintervention year that increased in intervention year 3 to a level equivalent to that in control sites.
from 4% to 12%. Kiefe et al\textsuperscript{31} used “achievable benchmarks of care,” that is, levels of performance achieved by top-performing centers. Those receiving the benchmark feedback improved delivery of influenza vaccine by 18%, compared with the control group. In contrast to these positive results, other trials seeking to improve delivery of β-receptor blocker use after myocardial infarction, clinical preventive services, compliance with national guidelines for the treatment of hypertension and depression, and compliance with protocols for the care of patients with AIDS failed to demonstrate significant changes in practice.\textsuperscript{32–35} A recently published meta-analysis of QI strategies for patients with diabetes mellitus demonstrated that most trials generated only modest improvements in glycemic control.\textsuperscript{4} The investigators also found evidence strongly suggesting publication bias, with smaller studies being more likely to show positive effects than larger studies.

Trials of QI in neonatology and pediatrics are more limited. Lozano et al\textsuperscript{36} reported the results of an intensive QI intervention conducted by Pediatric Asthma Care Patient Outcomes Research Task Force. A resource-intensive intervention using organizational change plus physician peer education was more effective than physician education alone.\textsuperscript{36} In neonatology, the Vermont Oxford Network used multifaceted QI techniques to improve patient outcomes focused on rates of nosocomial infections (5.5% reduction, compared with 1.6% for nonparticipants) and BPD (12.5% reduction from 43.5% to 31%, compared with 8.3% reduction for nonparticipants).\textsuperscript{21} The Vermont Oxford Network group reported a subsequent study that enrolled self-selected centers focused on BPD and showed reductions in rates of BPD in before/after comparisons.\textsuperscript{37} In another study, the Vermont Oxford Network investigators demonstrated outcomes similar to those of the current trial.\textsuperscript{38} Implementation of a multimodal QI intervention resulted in earlier administration of surfactant, compared with control centers, but those improved practices did not translate into improved patient outcomes, measured as death or pneumothorax.

One important difference between the current trial and previous studies of QI to reduce BPD rates was the outcome measure we used. Previous studies used a clinical definition of BPD defined by oxygen and/or ventilation exposure at 36 weeks, without controlling for oxygen saturation values delivered. As a prelude to this trial, we developed a rigorous definition of BPD that included a room air challenge for selected infants (those receiving <30% effective oxygen).\textsuperscript{33} This physiologic definition of BPD was applied equally in the intervention and control centers. As we reported previously, the definition resulted in a mean reduction of 10% in the rates of BPD (range: 0%–44%). The implementation of this definition could be considered an intervention that focused clinicians’ attention on the importance of integrating oxygen delivery, especially through a nasal cannula, with saturation monitoring at both intervention and control centers.\textsuperscript{39} It might have been the most effective intervention in the trial, dwarfing the effects of other potentially beneficial practices.

What accounts for our finding that the multimodal intervention failed to improve patient outcomes? One possibility is that, despite randomization, there were important random differences between intervention and control centers in the preexisting rates of survival free of BPD. However, the rates were comparable (63.3% in the intervention centers and 62.8% in the control centers in year 1). Another possibility is that our trial was not large enough to identify a clinically important benefit from benchmarking. The 95% CIs for the changes in rates of survival free of BPD resulting from benchmarking and QI excluded a benefit greater than a 4.4% improvement in the rate of survival without BPD and included a hazard as great as a 6.1% increase in the rate of death or BPD. These CIs indicate that the trial was large enough to exclude important larger effects.\textsuperscript{40,41} Another possibility is that the QI training was ineffective. We think that this was not the case, because intervention centers demonstrated greater practice changes than did control centers.

A final (and we think more likely) possibility is that adopting practices from centers with exemplary outcomes may not be beneficial when there is only weak evidence supporting these practices. Well-controlled studies reporting benefits from benchmarking largely have been studies promoting the use of interventions established previously as beneficial in randomized trials. The interventions in our study were those with the strongest available evidence. However, few (such as early administration of surfactant) have been shown to be beneficial in randomized trials. Of the myriad of practice differences between centers, it remains to be established whether the practices that result in superior outcomes in benchmark centers can be reliably recognized and implemented by visiting health care teams. It is even possible that some interventions selected in QI efforts affect outcomes adversely.\textsuperscript{42,43} An example of this is the selection of a skin emollient as an intervention to reduce infection by one Vermont Oxford Network collaborative group. Emollient was shown to reduce infection rates in a single-center study but was later shown to increase infection rates in a large randomized trial.\textsuperscript{44,45} Introducing change, no matter how well intended, may perturb a stable system, with potentially adverse outcomes. The apparent increase in rates of severe intraventricular hemorrhage among neonates with PMA of <26 weeks in intervention centers may be a statistical anomaly or a real but unintended adverse consequence of changes in care. It is possible that interventions in the delivery room prolonged the time spent in the delivery

---

886  WALSH et al
room and contributed to hypothermia and associated intraventricular hemorrhage.

In this cluster-randomized, controlled trial, NICU teams trained in benchmarking and QI techniques benefited from the intervention with practice changes but did not improve rates of survival free of BPD in neonates with birth weights of <1250 g, compared with centers continuing usual practice. These results have implications for the design of future QI trials, in that other interventions may be required to produce change. Additional refinements are needed to create and to maintain larger magnitudes of change and to improve patient outcomes.

APPENDIX: DESCRIPTION OF POTENTIALLY BETTER PRACTICES

Practices in the Delivery Room

1. Fellow or attending physician present at every high-risk delivery. Resuscitation of high-risk infants was led by a fellow or attending physician at the benchmark sites. Level of evidence: not available; metric: audit during site visits.

2. Respiratory therapist present at every high-risk delivery. The resuscitation team included a respiratory therapist at the benchmark sites. Level of evidence: not available; metric: audit during site visits.

3. Consistent equipment in all delivery rooms. Each resuscitation site was configured with identical equipment, to facilitate the resuscitation of high-risk patients. Level of evidence: not available; metric: audit during site visits.

4. Limited tidal volumes used in manual ventilation during resuscitation. At the benchmark sites, teams focused on limiting tidal volumes during resuscitation by assessing chest wall excursion visually or measuring delivered tidal volumes. The goal was to have barely visible chest wall movements. Evidence in animal models supports this concept. Level of evidence: not available; metric: audit during site visits.

5. Prophylactic use of surfactant. At 2 benchmark centers, infants at gestational ages of <28 weeks were immediately intubated and given surfactant. The third center emphasized CPAP beginning in the delivery room. Strong evidence supports this concept to minimize the severity of respiratory distress syndrome. Level of evidence: level 1A; metric: time to surfactant use in the delivery room.

6. Use of device to provide positive end-expiratory pressure and to limit tidal volume. This practice was not in place at the benchmark centers. Intervention teams added this to the list of potentially better practices to support limitation of delivered tidal volumes. Level of evidence: not available; metric: audit during site visits.

Respiratory Care Practices

7. Selective intubation with liberal use of CPAP. One benchmark center emphasized CPAP beginning in the delivery room. The other 2 centers used intubation with prophylactic surfactant treatment. The evidence for a primary CPAP strategy is weak, with reports from small case series and observational studies. Level of evidence: indeterminate; metric: proportion of infants treated with CPAP on admission to the NICU.

8. Early use of surfactant if intubated. The benchmark center using a primary CPAP strategy administered surfactant at once if a decision to intubate the infant was made. Evidence supports this concept to minimize the severity of respiratory distress syndrome. Level of evidence: level 1A; metric: time to surfactant administration in the NICU.

9. Assessment of volume/pressure and targeting of lowest levels to achieve modest chest rise and to avoid exuberant chest wall motion if intubated. All 3 benchmark centers focused on low tidal volumes for intubated infants. All 3 centers used pressure-controlled, time-cycled ventilators. None had the capacity to measure tidal volume, and instead they used physical examination and assessment of Pco2 to limit tidal volumes. Level of evidence: level 5B; metric: mean peak inspiratory pressure on days 1 and 3 for all intubated neonates.

10. Aggressive weaning and early extubation if intubated. Two benchmark centers weaned patients aggressively and extubated them without birth weight or postnatal age limitations. In these 2 centers, teams noted during the site visits that it was common to see tiny infants at 24 hours of age receiving CPAP. The third center did not extubate patients but weaned them in the first 24 hours to low ventilatory rates and tidal volumes that were comparable to those delivered with CPAP. Level of evidence: level 5; metric: duration of ventilation in the first 7 days of life.

11. Higher PaCO2 targets for all patients. The 3 benchmark centers accepted higher PaCO2 levels to permit weaning from ventilators and/or the use of CPAP. Experimental animal data and uncontrolled human observational studies support permissive hypercapnia as a protective strategy. Level of evidence: indeterminate; metric: mean PaCO2 for all patients with measurements on days 1 and 3 of life.

12. Lower oxygen saturation goals. The 3 benchmark centers accepted lower oxygen saturation targets of 85% to 90%. In addition, caregivers were noted, during the site visits, to observe infants during desaturation events without increasing oxygen supplementation, with the goal of allowing the infant to resolve the desaturation event independently. Level of evidence: indeterminate; metric: mean PaO2 measured at 4 time points daily during the first 7 days of life for all those with measurements.

13. High-saturation alarm set at 95%. The 3 bench-
mark centers set the oxygen saturation alarm at 95% and rapidly weaned patients from supplemental oxygen when the saturation range exceeded the target. Level of evidence: not available; metric: audit during site visits.

14. Avoidance of routine suctioning for patients undergoing ventilation. The 3 benchmark centers avoided suctioning for patients undergoing ventilation that was set by a time schedule and instead assessed the patients at intervals and suctioned as needed. In addition, 2 centers used an inline suction system. Level of evidence: indeterminate; metric: audit during site visits.

15. Avoidance of hand-bagging for patients undergoing ventilation. One benchmark center prohibited the practice of ventilating with an anesthetic or self-inflating bag, as a method to limit exposure to unregulated tidal volumes. Level of evidence: not available; metric: audit during site visits.

16. Nonroutine use of analgesics/sedatives for patients undergoing ventilation. None of the benchmark centers routinely administered analgesics/sedatives to patients treated with mechanical ventilation. Instead, comfort techniques such as swaddling were used. Level of evidence: indeterminate; metric: audit during site visits.

17. Prophylactic use of methylxanthines before extubation. Routine administration of methylxanthines before extubation was not used at the benchmark centers. Level of evidence: level 2; metric: not selected.

18. Consensus regarding ventilatory management. At 2 benchmark centers, there was high consistency in ventilator management practices among individual physicians and teams. Level of evidence: level 2; metric: not selected.

19. Limited intravenous fluids. At all 3 benchmark centers, intravenous fluids were initiated at 80 to 100 mL/kg per day and were adjusted by using daily weight goals. The intent was for weight loss to occur in the first 7 days of life. Level of evidence: level 2; metrics: mean intravenous fluid intake on days 1 and 3 and percentages of weight loss on days 3 and 7.

20. High-humidity environments. Two benchmark centers used high-humidity environments to limit intravenous fluid administration. Level of evidence: level 5; metric: audit during site visits.

21. Limited volume expansion to treat low blood pressure. One benchmark center used protocols to decrease treatment of low blood pressure, to limit intravenous fluid administration. Level of evidence: level 5; metric: not selected.

22. Aggressive approach to patent ductus arteriosus. One benchmark center used prophylactic indomethacin treatment, and 1 center screened for patent ductus arteriosus and treated patients with indomethacin in the first 24 hours. The third center had high rates of patent ductus arteriosus ligation. The patent ductus arteriosus was ligated for patients who experienced failure of indomethacin treatment and had persistent oxygen requirements exceeding 40%. Level of evidence: level 5; metric: not selected.

23. Early introduction of parenteral protein intake. Two benchmark centers began total parenteral nutrition administration on admission to the NICU. Level of evidence: level 5; metric: not selected.


25. Full total parenteral nutrition with increasing enteral feeding. Two benchmark centers maintained total parenteral nutrition at 100 to 120 mL/kg per day as enteral nutrition was increased. Level of evidence: not available; metric: not selected.

26. Frequent use of human milk. Two benchmark centers promoted the use of human milk and had programs in place to support human milk feedings. Both centers had human milk administered to >60% of their patients. Level of evidence: indeterminate; metric: not selected.

27. Vitamin A prophylaxis. One benchmark center used vitamin A prophylaxis. The other 2 centers chose not to implement prophylaxis because their BPD rates were low and nurses objected to intramuscular injections. Level of evidence: level 2; metric: audit during site visits.

Fluid and Nutrition Practices

19. Limited intravenous fluids. At all 3 benchmark centers, intravenous fluids were initiated at 80 to 100 mL/kg per day and were adjusted by using daily weight goals. The intent was for weight loss to occur in the first 7 days of life. Level of evidence: level 2; metrics: mean intravenous fluid intake on days 1 and 3 and percentages of weight loss on days 3 and 7.

20. High-humidity environments. Two benchmark centers used high-humidity environments to limit intravenous fluid administration. Level of evidence: level 5; metric: audit during site visits.

21. Limited volume expansion to treat low blood pressure. One benchmark center used protocols to decrease treatment of low blood pressure, to limit intravenous fluid administration. Level of evidence: level 5; metric: not selected.

22. Aggressive approach to patent ductus arteriosus. One benchmark center used prophylactic indomethacin treatment, and 1 center screened for patent ductus arteriosus and treated patients with indomethacin in the first 24 hours. The third center had high rates of patent ductus arteriosus ligation. The patent ductus arteriosus was ligated for patients who experienced failure of indomethacin treatment and had persistent oxygen requirements exceeding 40%. Level of evidence: level 5; metric: not selected.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Child Health and Human Development (grants U10 HD34216, U10 HD21364, U10 HD27853, U10 HD27851, U01 HD36790, U10 HD27856, U10 HD21397, U10 HD27881, U10 HD27880, U10 HD21415, U10 HD21373, U10 HD21385, U10 HD27871, U10 HD34167, and U10 HD27904) and the General Clinical Research Center (grants M01 RR08084, M01 RR00750, M01 RR00997, M01 RR00070, M01 RR06022, M01 RR02635, M01 RR02172, and M01 RR01032). The funding agency provided overall oversight for study conduct. All data analyses and interpretation were independent of the funding agency.

Study participants were as follows: advisory committee: M. C. Walsh, MD, MS, A. A. Fanaroff, MD, Case Western Reserve University (Cleveland, OH); A. H. Jobe, MD, PhD, University of Cincinnati (Cincinnati, OH); R. Higgins, MD, National Institute of Child Health and Human Development (Bethesda, MD); N. Finer, MD, University of California, San Diego (San Diego, CA); K. Poole, PhD, Research Triangle Institute (Research Triangle Park, NC); training committee: Duncan Neuhauser,
REFERENCES

1. Institute of Medicine, Committee on Quality of Health Care in America. To Err Is Human: Building a Safer Health System. Washington, DC: National Academies Press; 2000


26. Linder W, Bosveck S, Hummler H, Pohlandt F. Delivery room

PhD, Case Western Reserve University (Cleveland, OH); Leslie Clarke, RN, MSN, MBA, Rainbow Babies & Children’s Hospital (Cleveland, OH); Lynn Lostocco, RN, MSN, National Association of Children's Hospitals (Warwick, RI); Neil Finer, MD, University of California, San Diego (San Diego, CA); intervention centers: S. Nada Kazzi, MD, MPH, K. Hayes-Hart, RN, M. Betts, RRT, S. Shankaran, MD, G. Muran, RN, Wayne State University (Detroit, MI); A. Laptook, MD, M. Martin, RN, J. Allen, RRT, University of Texas Southwestern (Dallas, TX); W. A. Engle, MD, L. Miller, RN, R. Hooper, RRT, J. Lemons, MD, Indiana University (Indianapolis, IN); W. Rhine, MD, C. Kibler, RN, J. Parker, RRT, D. Stevenson, MD, B. Ball, BS, Stanford University (Palo Alto, CA); M. Rasmussen, MD, M. Grabarczyk, BSN, C. Joseph, RRT, K. Arnell, BSN, Sharp Mary Birch Hospital for Women (San Diego, CA); G. Heldt, MD, R. Bridge, RN, J. Good- mar, RRT, N. Finer, MD, C. Henderson, RCP, University of California, San Diego (San Diego, CA); S. Buchter, MD, M. Berry, RN, I. Seabrook, RRT, B. Stoll, MD, E. Hale, RN, Emory University (Atlanta, GA); benchmark centers: S. Duara, MD, R. Everette, RN, University of Miami (Miami, FL); W. Carlo, MD, M. Collins, RN, Univer-

sity of Alabama (Birmingham, AL); W. Oh, MD, A. Hensman, RN, Brown University (Providence, RI); control centers: M. T. O’Shea, MD, MPH, N. Peters, RN, Wake Forest University (Winston-Salem, NC); J. Tyson, MD, MPH, G. McDavid, RN, University of Texas (Hous-

ton, TX); A. A. Fanaroff, MD, M. C. Walsh, MD, MS, N. Newman, RN, Case Western Reserve University (Cleve-

land, OH); D. Phelps, MD, L. Reubens, RN, University of Rochester (Rochester, NY); R. A. Ehrenkranz, MD, P. Gettner, RN, Yale University (New Haven, CT); C. Mi-

chael Cotten, MD, K. Auten, RN, Duke University (Durham, NC); E. Donovan, MD, C. Grisby RN, Univer-

sity of Cincinnati (Cincinnati, OH); statistical center: Qing Yao, PhD, Ken Poole, PhD, Research Triangle In-

stitute (Research Triangle Park, NC).

We thank the nursing and medical staff members and parents of the patients in the units for their diligent implementation of this complex trial. We also thank the Neonatal Research Network Research coordinators and study nurses, without whom the trial could not have been completed.


41. Detsky AS, Sackett DL. When was a “negative” clinical trial big enough? How many patients you needed depends on what you found. *Arch Intern Med.* 1985;145:709–712

42. Wachter RM. Expected and unanticipated consequences of the quality and information technology revolutions. *JAMA.* 2006;295:2780–2783


How Reliable Is a Negative Blood Culture Result? Volume of Blood Submitted for Culture in Routine Practice in a Children’s Hospital

Thomas G. Connell, MRCPIa,b,c, Mhisti Rele, MDd, Donna Cowley, BAppSci4, Jim P. Buttery, FRACPa,b,c, Nigel Curtis, FRCPCH, PhD4a,b,c

aInfectious Diseases Unit, Department of General Medicine, bMurdoch Children’s Research Institute, and cDepartment of Microbiology, Royal Children’s Hospital Melbourne, Parkville, Australia; dDepartment of Paediatrics, University of Melbourne, Parkville, Australia

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

ABSTRACT

OBJECTIVES. The primary aims of this study were to determine the volume of blood submitted for culture in routine clinical practice and to establish the proportion of blood cultures with a blood volume inadequate for reliable detection of bacteremia.

METHODS. The volumes of blood samples submitted for culture from infants and children up to 18 years of age were measured over a 6-month period. Blood cultures were deemed adequate submissions if they contained an appropriate (age-related) volume of blood and were submitted in the correct blood culture bottle type. During the study, an educational intervention designed to increase the proportion of adequate blood culture submissions was undertaken.

RESULTS. The volume of blood submitted in 1358 blood culture bottles from 783 patients was analyzed. Of the 1067 preintervention blood cultures, 491 (46.0%) contained an adequate blood volume and only 378 (35.4%) were adequate submissions on the basis of collection into the correct blood culture bottle type. After the intervention, there were significant increases in both the proportion of blood cultures containing an adequate blood volume (186 [63.9%] of 291 cultures) and the proportion of adequate submissions (149 [51.2%] of 291 cultures). Overall, blood cultures with an adequate blood volume were more likely than those with an inadequate blood volume to yield positive blood culture results (34 [5.2%] of 655 cultures vs 14 [2.1%] of 648 cultures). Similarly, adequate blood culture submissions were more likely than inadequate submissions to yield positive blood culture results (26 [5.1%] of 506 cultures vs 22 [2.8%] of 797 cultures).

CONCLUSIONS. In routine clinical practice, a negative blood culture result is almost inevitable for a large proportion of blood cultures because of the submission of an inadequate volume of blood. Even after an educational intervention, nearly one half of blood cultures were inadequate submissions.
Blood cultures remain the standard method for detecting bacteremia in the evaluation of sick infants and children. The isolation of an organism confers many advantages, including the optimal choice and duration of antibiotic treatment. The exclusion of bacteremia is also important, because it may enable the cessation of antibiotic treatment and consequently reduce the length and cost of hospital stay, as well as decreasing the development of antimicrobial resistance. The ability to exclude bacteremia on the basis of a negative blood culture result depends on the sensitivity and negative predictive value of this test. Many factors influence the yield from blood cultures but the single most important factor is blood volume. Evidence from both adult and pediatric studies shows that the rate of isolation from blood cultures increases with the quantity of blood submitted. The interpretation of negative blood culture results is based on studies using adequate volumes of blood in research settings. When the volume of blood submitted for culture is inadequate, a negative blood culture result is potentially misleading in falsely excluding significant bacteremia. The aim of this study was to determine the volume of blood inoculated into blood culture bottles in a tertiary children’s hospital in routine clinical practice and to determine the proportion of blood cultures with a blood volume inadequate for reliable detection of bacteremia. The effect of an educational intervention designed to increase the rate of adequate blood culture submissions was also investigated.

METHODS
All blood cultures from infants and children up to 18 years of age that were received in the microbiology laboratory at the Royal Children’s Hospital Melbourne during a 6-month period were included in the study. Blood cultures were drawn almost exclusively by medical or nursing (central line only) staff members. To ensure that the results represented routine clinical practice, these staff members were not informed of the study. Blood culture bottles were weighed with an Ohaus TS400D top-loading balance, coded numerically, and distributed to clinical areas for use in the normal manner. The bottles included “pediatric” (yellow, 20 mL, BacT/Alert PF/Pediatric; Organon-Teknika, Durham, NC) and “nonpediatric” (green, 30 mL, BacT/Alert FA; Organon-Teknika) bottles used for the detection of aerobic and facultative anaerobic organisms in an automated blood culture system (BacT/Alert; Organon-Teknika). The manufacturers recommend the use of pediatric bottles for inoculation of blood volumes up to 4 mL and nonpediatric bottles for volumes from 5 to 10 mL. Anaerobic blood culture bottles were not included in the study because they represent only a small proportion (<1%) of all blood cultures submitted to the laboratory.

On arrival in the microbiology laboratory, the bottles were reweighed before subsequent normal processing for culture. The final blood culture weight was determined by subtracting the initial bottle weight (adjusted for the weight of the cap) from the bottle weight on receipt in the laboratory, with an additional adjustment for weight loss attributable to storage duration before use. To determine the weight variation over time under normal storage conditions, 20 capped blood culture bottles stored at room temperature, away from direct sunlight, were weighed at regular intervals during the study period. The final blood volume was determined by dividing the final blood weight by a factor of 1.06.

The following details were recorded for each sample: date of blood culture, patient’s date of birth, and site of blood culture. The site of blood collection was classified as central or unspecified. The age of the patient at the time of blood culture was determined.

In this study, an adequate blood culture volume was defined as ≥0.5 mL for patients <1 month of age, ≥1.0 mL for patients between 1 month and 36 months of age, and ≥4.0 mL for patients ≥36 months of age. An adequate blood culture submission was deemed to have occurred when an adequate blood volume was collected into the correct bottle type. The use of a pediatric bottle was defined as correct if the volume of blood was <5.0 mL and a nonpediatric bottle if the volume was ≥4.0 mL. The use of either bottle for volumes between 4.0 and 5.0 mL was deliberately deemed correct, to take into account the fact that estimation of collected blood volume at the point of sampling is not precise.

An intervention designed to encourage increased awareness among clinical staff members about appropriate blood culture volumes and bottle types was undertaken during the study. The intervention consisted of the introduction in all blood collection areas of laminated posters indicating the blood volumes recommended for different ages, with guidance on correct bottle use.

Because the overall rate of positive blood cultures at our institution is low, the study was not designed to assess the influence of blood culture volume on subsequent isolation rates. We did, however, analyze the blood culture results to determine whether blood cultures with an adequate blood volume or an adequate submission were more likely to yield positive bacterial culture results. For this analysis, coagulase-negative staphylococci (except from neonates), diphtheroid species, Micrococcus spp, and nonpneumococcal α-hemolytic streptococci were classified as contaminants. Statistical significance testing was undertaking by using the χ² test or Fisher’s exact test for categorical data and the Mann-Whitney test for continuous data.
RESULTS

During the study period, 1358 blood culture bottles collected from 783 patients were included in the study. Of those patients, 618 (78.9%) had a single blood culture performed, with the rest having 2 (n = 77; 9.9%), 3 (n = 23; 2.9%), or ≥4 (n = 65; 8.3%) blood cultures. In almost all instances, these represented blood cultures undertaken on different occasions, because it is normal practice in our institution to inoculate only a single blood culture bottle. The median patient age was 3.9 years (range: 0.0–18.7 years). Of 1358 blood cultures, 133 (9.8%) were collected from patients <1 month of age, 469 (34.5%) from patients between 1 and 36 months of age, and 756 (55.7%) from patients ≥36 months of age. Pediatric blood culture bottles were used for 859 blood cultures (63.3%) and nonpediatric bottles for the rest.

Of the 1358 blood cultures, 1067 (78.6%) were collected before the intervention and the rest after. Although there was no significant difference in age distribution between the 2 groups overall (P = .06, Mann-Whitney test), there were small but significant differences between the preintervention and postintervention groups in the proportions falling within 2 of the 3 age categories, with a smaller proportion of blood cultures from patients in the <1-month group (8.9% vs 13.1%; P = .04, Fisher’s exact test) and a correspondingly larger proportion of blood cultures from patients in the ≥36-month group (57.3% vs 49.8%; P = .03) in the preintervention group.

The blood volumes submitted for culture in each age category are shown in Table 1 and Fig 1. Of the 1067 preintervention blood cultures, 491 (46.0%) contained an adequate blood volume. Of those 491 blood cultures, 378 (77.0%) were deemed adequate blood culture submissions on the basis of collection into the correct blood culture bottle. Therefore, overall 378 (35.4%) of 1067 submissions were classified as adequate blood culture submissions (Table 1).

After the intervention, there were increases in all age groups in both the proportion of blood cultures with adequate blood volume and the proportion of these collected into the correct blood culture bottle. Specifically, of the 291 blood cultures taken after the intervention, 186 (63.9%) contained an adequate blood volume, of which 149 (80.1%) were collected into the correct blood culture bottle. Therefore, overall 149 (51.2%) of 291 blood cultures were considered adequate blood culture submissions after the intervention (P < .0001).

Culture results were available for 1303 (95.9%) of the 1358 blood cultures in the study (1021 [95.7%] of 1067 preintervention cultures and 282 [96.9%] of 291 postintervention cultures; P = .3). During the study period, bacteria were cultured from 96 (7.4%) of these 1303 blood cultures. Of these 96 blood cultures, 48 (50.0%) grew likely contaminants and the rest grew potentially pathogenic organisms, meaning that there were 48 (3.7%) true-positive blood cultures overall.

Blood cultures with an adequate blood volume were more likely than those with an inadequate blood volume to yield positive blood culture results (34 [5.2%] of 655 cultures vs 14 [2.2%] of 648 cultures; P = .005). Similarly, adequate blood culture submissions were more likely than inadequate submissions to yield noncontaminant positive blood culture results (26 [5.1%] of 506 cultures vs 22 [2.8%] of 797 cultures; P = .03). The increased likelihood of a positive culture result with an adequate blood volume remained significant even when coagulase-negative staphylococci were excluded from the analysis (29 [4.4%] of 655 cultures vs 13 [2.0%] of 648 cultures; P = .02). Of the 1358 blood cultures overall, 169 (12.4%) were submitted with a volume of <0.5 mL. Notably, within this group, a pathogenic organism (Enterobacter cloacae) was cultured from only 1 blood culture.

Of the 859 blood cultures submitted in a pediatric blood culture bottle, 77 (8.9%) were deemed inadequate because the volume exceeded the recommended maximum of 5 mL. These were defined as inadequate because an excess volume of blood, relative to broth, may diminish the sensitivity of blood culture.22 Reanalysis of the data with these 77 blood cultures categorized as inadequate did not alter the main findings of our study. In this alternative analysis, 418 (39.2%) of the 1067 preintervention blood cultures were adequate submissions and there was still a significant increase in the proportion of inadequate submissions after intervention.

| Table 1 Volumes Submitted for 1358 Blood Cultures According to Age Category |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Age Group | Before Intervention | After Intervention |
|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|           | Final Blood Volume, Median (Range), mL | Adequate Volume, n/N (%) | Adequate Submission, n/N (%) | Final Blood Volume, Median (Range), mL | Adequate Volume, n/N (%) | Adequate Submission, n/N (%) |
| <1 mo     | 0.68 (0.0–2.6) | 62/95 (65.2) | 43/62 (69.3) | 0.85 (0.2–3.4) | 31/38 (81.6) | 31/31 (100) |
| 1 < 36 mo | 1.24 (0.0–16.6) | 218/361 (60.4) | 159/218 (72.9) | 1.59 (0.1–11.6) | 80/108 (74.1) | 60/80 (75.0) |
| ≥36 mo    | 2.49 (0.0–20.1) | 211/611 (34.5) | 176/211 (83.4) | 4.27 (0.0–11.6) | 75/145 (51.7) | 58/75 (77.3) |
| All ages  | 1.75 (0.0–20.1) | 491/1067 (46.0) | 378/491 (77.0) | 2.12 (0.2–20.6) | 186/291 (63.9) | 149/186 (80.1) |

a P ≤ .01 for difference between preintervention and postintervention volumes (Mann-Whitney test) and proportions of adequate volumes or submissions (Fisher’s exact test).
b P ≤ .005.
c P ≤ .0001.
adequate submissions after the intervention (178 [61.2%] of 291 cultures; *P* < .0001). In addition, adequate blood culture submissions were still more likely overall than inadequate submissions to yield positive blood culture results (29 [5.1%] of 574 cultures vs 19 [2.6%] of 729 cultures; *P* < .001). Interestingly, under normal storage conditions, there was a noteworthy loss of weight in the control, capped, blood culture bottles. The mean ± SD weight loss from the 20 bottles during the study period was 0.72 ± 0.0100 g. This weight loss was taken into account, on a pro rata basis, when the final blood volumes were calculated.

**DISCUSSION**

The finding of a negative blood culture frequently influences management. Our finding that, in routine clinical practice at a tertiary children’s hospital, over half of blood cultures contained a volume of blood that was inadequate to enable a negative result to exclude bacteremia reliably has important implications. A negative result is almost invariably interpreted without regard for the volume of blood that was submitted and therefore without a true appreciation of the test’s sensitivity or negative predictive value in a given patient. In many instances in our study, the blood culture submitted did not just constitute a test with decreased sensitivity but was equivalent to no meaningful test having occurred, because the blood volume submitted was too small to have a reasonable chance of leading to the detection of even high-level bacteremia. Specifically, of 1358 blood cultures, 169 (12.4%) were submitted with <0.5 mL of blood, and this proportion increased to 40 (30%) of 133 cultures for patients <1 month of age. It is worrying that a negative blood culture result is potentially misleading for up to one third of neonatal cultures. Although our study was not designed to assess the influence of blood volume on subsequent isolation rates, blood cultures submitted with an adequate blood volume were more likely to yield a positive bacterial culture than were those with an inadequate volume. Conversely, a positive culture result was obtained for only 1 of the 169 blood cultures submitted with <0.5 mL of blood.

For small infants and children, phlebotomy can be difficult and in certain situations potentially harmful. Although smaller-volume blood samples are therefore submitted from children, this is offset to an extent by the fact that the level of bacteremia is usually higher in infants and young children. In one study, the average

![Figure 1](image_url)

**FIGURE 1**
Volumes submitted for 1358 (1067 preintervention and 291 postintervention) blood cultures. The bars represent the median volume for each age group, and the boxes encompass blood cultures submitted with an adequate volume for the patient’s age. *P* ≤ 0.01; †*P* ≤ 0.001; and ‡*P* ≤ 0.001 for differences in median volume before intervention and after intervention in each age group.
bacterial count for children with *Haemophilus influenzae* bacteremia was 6000 colony-forming units (CFUs) per mL of blood and that for children with *Streptococcus pneumoniae* was 50 CFUs per mL of blood. In another study, 60% of children with significant isolates had bacterial counts of >10 CFUs per mL. For a series of 30 newborns with *Escherichia coli* bacteremia, a high bacterial count (>1000 CFUs per mL) was found in 11 cultures (31%). In that study, however, a significant proportion (23%) of cultures had colony counts between 0 and 4 CFUs per mL, which suggests that the level of bacteremia in small infants and children may sometimes be low and therefore more difficult to detect in small-volume blood cultures. A large study of infants from birth to 2 months of age showed that two thirds of isolates detected had colony counts of <10 CFUs per mL. Furthermore, in an in vitro study, 0.5 mL of adult blood spiked with pediatric pathogens was inadequate for the detection of low-level bacteremia (<4 CFUs per mL).

Currently, there are limited data about the optimal volume for blood culture in children. Much of the available information has been extrapolated from adult studies, from which it is clear that isolation rates increase proportionally with the volume of blood submitted for culture, with optimal culture volumes of 10 to 30 mL being recommended. In one study, 17% more clinically significant isolates were detected when 13 to 16 mL of blood was cultured, in contrast to 6.5 to 8 mL. Several studies have confirmed the increase in detection rates that results from increased blood culture volume, reporting figures of 0.6% to 4.7% increased yield for each extra 1 mL of blood cultured. In the pediatric population, a study of 300 children attending an emergency department reported an increased yield with a single 6-mL blood culture, compared with 2 separate 2-mL cultures. In that study, up to 10 mL of blood was drawn from each individual patient, allowing a within-patient comparison of different blood volumes. In a recent study of Kenyan children, there was an increase in the blood culture isolation rate with increasing volume, from 5.6% at 1 mL to 6.8% at 2 mL and 7.9% at 3 mL. In a study of immunocompromised children, the introduction of a policy to increase blood culture volume led to an increase in the number of significant isolates recovered. In contrast, there was no difference in overall isolation rates between 0.5-mL and 1.5-mL blood cultures (13.4% vs 13.1%) in a recent study that investigated 2 blood culture procedures in a PICU. However, coagulase-negative staphylococci were isolated more commonly with the low-volume blood culture system, and it was not possible to determine whether they were significant or contaminant isolates.

Despite these studies, it is difficult to establish the optimal volume to detect circulating bacteria in infants and children. Although few definitions of an “adequate” blood culture volume exist, minimal volumes of 0.5 mL to 1 mL for infants and 1 to ≥30 mL for older children have been recommended. Our definition of adequate volume was deliberately designed to be conservative, thereby giving a minimal estimate of the rate of inadequate blood cultures. A more-stringent definition requiring a minimum of 1 mL for children <36 months of age and 6 mL for older children would have increased the proportion of blood cultures with inadequate volume before the intervention to 719 (67.3%) of 1067 cultures overall (70 [73.6%] of 95 cultures for the group of patients <1 month of age and 506 [82.8%] of 611 cultures for patients ≥36 months of age).

Our study showed that, even when the blood culture volume is adequate, the sensitivity of blood cultures is reduced by the frequent use of an incorrect bottle type. Blood culture bottles are designed for the incubation of a specific range of blood volumes. The inoculation of inappropriately large or small amounts of blood into blood culture bottles may result in decreased isolation rates as a result of altered blood/broth ratios. In addition, we showed that a relatively simple and reproducible intervention to educate staff members about the importance of the correct collection of blood cultures was associated with an increase in the proportion of adequate blood culture submissions.

The gradual decrease in weight of capped blood culture bottles during storage under normal conditions has not been reported previously. Evaporation of medium through the cap is the most likely explanation. By taking this weight loss into account when calculating final blood volumes, we avoided underestimating the volumes collected.

Possible limitations to our study include the inability to determine the clinical indications for the blood cultures, because frequently clinical details were not provided on request forms. Therefore, we were unable to ascertain which patients were thought to be at high risk for bacteremia or had clinical sepsis. However, several studies have shown that almost 75% of blood cultures taken from patients admitted to the ICU with a clinical diagnosis of sepsis fail to recover an organism. In addition, although we found an association between adequate submission and isolation rates, our study was not designed primarily to determine the influence of blood culture volumes on subsequent isolation rates, because of our relatively small sample size.

Despite these limitations, our study highlights the fact that blood cultures are frequently performed inadequately. We believe that the results of our study are likely to be reflected at other institutions. To our knowledge, this is the first study in a pediatric population investigating the volume of blood submitted for culture in routine practice outside a study setting. It is worrying that, for a large proportion of blood cultures, a negative result is almost inevitable because of inadequate blood volume.
ACKNOWLEDGMENTS

We thank Dr Andrew Daley and the microbiology laboratory staff members, including Chris Pearce, for their cooperation and contribution to this study.

REFERENCES

Effectiveness of Trimethoprim/Sulfamethoxazole for Children With Chronic Active Otitis Media: A Randomized, Placebo-Controlled Trial

Erwin L. van der Veen, MD, Maroeska M. Rovers, PhD, Frans W. J. Albers, MD, PhD, Elisabeth A. M. Sanders, MD, PhD, Anne G. M. Schilder, MD, PhD

Departments of Otorhinolaryngology and Paediatric Immunology, Wilhelmina Children’s Hospital, and Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands

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

ABSTRACT

OBJECTIVE. The goal was to determine the clinical effectiveness of prolonged outpatient treatment with trimethoprim/sulfamethoxazole for children with chronic active otitis media.

METHODS. We performed a randomized, placebo-controlled trial with 101 children (1–12 years of age) with chronic active otitis media (defined as otorrhea for ≥12 weeks). In addition to a short course of steroid and antibiotic eardrops, children were assigned randomly to receive 6 to 12 weeks of orally administered trimethoprim/sulfamethoxazole (18 mg/kg, 2 times per day) or placebo and were monitored for 1 year.

RESULTS. At 6 weeks, 28% of children in the trimethoprim/sulfamethoxazole group and 53% of children in the placebo group had otomicroscopic signs of otorrhea. At 12 weeks, these values were 32% and 47%, respectively. At 1 year, the numbers of children with otorrhea were similar in the 2 groups (25% and 20%, respectively). One child in the trimethoprim/sulfamethoxazole group developed a skin rash. Vomiting or diarrhea was reported for 9% of the trimethoprim/sulfamethoxazole group and 2% of the placebo group. Pure-tone hearing levels and health-related quality of life improved during the study but did not differ between the trimethoprim/sulfamethoxazole group and the placebo group. Pseudomonas aeruginosa was the most frequently isolated bacteria in the otorrhea samples from both groups.

CONCLUSIONS. A 6- to 12-week course of high-dose, orally administered trimethoprim/sulfamethoxazole therapy is beneficial for children with chronic active otitis media. The treatment effect is most pronounced with the shorter course and disappears if administration of the medication is discontinued.
Chronic active (mucosal) otitis media (COM) is a common infectious disease, affecting children in both developing and industrialized countries. It causes considerable morbidity and is a major global cause of hearing impairment in children. Moreover, it may lead to serious extracranial and intracranial complications, such as mastoiditis and meningitis. An active approach to serious extracranial and intracranial complications, hearing impairment in children. Moreover, it may lead to considerable morbidity and is a major global cause of morbidity. It is well-known that the management of COM is therefore important.

Evidence regarding the most effective medical or surgical treatment of COM is incomplete; few randomized, controlled trials have been performed, and inclusion criteria, outcome measures, and follow-up methods in those studies vary considerably.

Trimethoprim/sulfamethoxazole has been used for many years for the treatment of upper respiratory tract infections in children. It is an inexpensive antimicrobial drug and is well tolerated by children. When used for prophylaxis for recurrent acute otitis media, it was found to be effective.

A retrospective analysis of data for children with COM treated with trimethoprim/sulfamethoxazole for a prolonged period at our hospital showed promising results. We therefore initiated a randomized, placebo-controlled trial of a 6- to 12-week course of orally administered trimethoprim/sulfamethoxazole in addition to a short course of steroid and antibiotic ear-drops for children with COM who had experienced failure of conventional management with topical medications and/or short-term systemic antibiotic treatment. In this article, we report on both the clinical effectiveness and bacteriologic findings.

Methods

Patients

We conducted a randomized, placebo-controlled trial between February 2003 and June 2006. Otorhinolaryngologists and pediatricians from across the Netherlands referred potential participants (ie, children with COM who had experienced failure of conventional management with topical medications and/or short-term systemic antibiotic treatment) to the pediatric otorhinolaryngology department of the University Medical Center Utrecht. Inclusion criteria were age of 1 to 12 years and a documented history of ≥3 months of continuous otorrhea through either a tympanic membrane perforation or a tympanostomy tube. We excluded children with (1) cholesteatoma, (2) known immunodeficiency other than for IgA or IgG subclasses, (3) Down syndrome, (4) craniofacial anomalies, (5) cystic fibrosis, (6) primary ciliary dyskinesia, (7) allergy to trimethoprim/sulfamethoxazole, or (8) continuous use of antibiotics for >6 weeks in the past 6 months. The medical ethics committee of the University Medical Center Utrecht approved the study protocol.

Randomization

Children whose parents gave informed consent were assigned randomly to receive either trimethoprim/sulfamethoxazole (18 mg/kg, administered orally, 2 times per day) or placebo for 6 to 12 weeks. A computer-generated randomization list was prepared by an independent data manager and sent to the hospital pharmacist, who then provided numbered boxes with bottles filled with a blinded suspension of either trimethoprim/sulfamethoxazole or placebo, with identical taste, bottle appearance, and fluid appearance. At entry into the trial, the investigator responsible for seeing the study participants allocated the next available number on the randomization list and the corresponding box with blinded suspension to each participant. The investigators remained blinded to the randomization until the end of the study.

When otorrhea was found to be present in either ear at the first control visit after 6 weeks, study medication administration was continued for another 6 weeks. Administration of the study medication was discontinued if both ears were found to be free of otorrhea and parents confirmed that they had seen no signs of otorrhea during the previous week. Parents were instructed to start study medication treatment again if symptoms of otorrhea recurred between the follow-up visits at 6 and 12 weeks. At inclusion and if otorrhea was present at the 6-week and 12-week follow-up visits, hydrocortisone/bacitracin/colistin eardrops were prescribed in addition to the study medication for 7 to 10 days, in both the trimethoprim/sulfamethoxazole group and the placebo group. These eardrops were chosen because they are widely used in the Netherlands and are generally considered safe. From July 2004 onward, hydrocortisone/neomycin/polymyxin B eardrops were prescribed because the former combination was no longer available in the Netherlands. During the first 12 weeks, local otorhinolaryngologists and pediatricians were allowed to prescribe additional medication, except for trimethoprim/sulfamethoxazole, to the participants according to their regular practice. After the second control visit at 12 weeks, administration of the study medication was discontinued irrespective of the presence or absence of otorrhea. At that time, children were referred back to their local doctors. Treatment was unblinded for an independent doctor, who informed the parents and local doctors about the assigned treatment by letter. The letter also included a treatment recommendation in case otorrhea was still present or recurred, that is, trimethoprim/sulfamethoxazole (18 mg/kg, 2 times per day, for 6–12 weeks) for the placebo group and azithromycin (5 mg/kg, once per day, for 6–12 weeks) for the trimethoprim/sulfamethoxazole group. At that stage, local otorhinolaryngologists and pediatricians were free to follow the recommendations or to manage symptoms of otorrhea according to their regular practice.
Inclusion and Follow-up Monitoring

At inclusion, disease-specific questionnaires, including information on potential risk factors for ear disease, duration of otorrhea before study entry, and previous treatments, were completed. At inclusion and at the 3 follow-up visits at 6 weeks, 12 weeks, and 1 year, parents completed 1 generic and 2 disease-specific questionnaires on health-related quality of life, namely, the Child Health Questionnaire parental form, a 6-item otitis media questionnaire, and a visual analog scale measuring ear-related quality of life. At these visits, the ears of the children were examined with an otomicroscope. The following features were noted: tympanostomy tube, tympanic membrane perforation, otorrhea, and middle ear effusion. If otorrhea was present, then a swab was taken from the otorrhea before suction cleaning was performed. To test for adverse reactions to the study medication, venous samples were taken for complete blood counts and hepatic (aspartate aminotransferase, alanine aminotransferase, and \(\gamma\)-glutamyltransferase) and renal (urea and creatinine) function tests at inclusion, at 6 weeks, and at 12 weeks. For children >3 years of age, pure-tone air conductive hearing levels were measured at frequencies of 500, 1000, 2000, and 4000 Hz.

Parents kept a diary of study medication and additional medication used for their child’s ear disease, including eardrops. These data were collected at the follow-up visits. During those visits, adverse effects of the study medication were noted and the empty and full bottles of study medication were weighed, to determine compliance rates.

Microbiologic Investigation

At inclusion and at the follow-up visits, study physicians took otorrhea samples by using flexible, sterile, rayon-tipped swabs (Medical Wire & Equipment Co, Corsham, Wiltshire, United Kingdom). The samples were immediately stored in Stuart’s transport medium at room temperature. Samples were transported to the microbiology laboratory and plated, within 18 hours after sampling, onto sheep blood (5%), \(Haemophilus\), and MacConkey agar plates for the isolation of potential aerobic pathogens. The culture plates were incubated aerobically at 37°C (MacConkey agar) and <5% carbon dioxide (blood and \(Haemophilus\) agars). They were examined at 24 and 48 hours. Colonies suspected to be Streptococcus pneumoniae, \(Haemophilus influenzae\), Moraxella catarrhalis, Staphylococcus aureus, \(Streptococcus pyogenes\), or aerobic Gram-negative bacteria were identified with previously described methods.
Study Outcomes

The primary end point was otomicroscopic signs of otorrhea in the presence of a tympanostomy tube or tympanic membrane perforation at follow-up times of 6 weeks, 12 weeks, and 1 year. Secondary outcome measures were use of medication other than the study medication for ear disease, adverse effects of the study medication, ear/nose/throat operations, health-related quality of life, pure-tone hearing levels, and bacteriologic findings for the otorrhea samples.

Statistical Aspects

Assuming a spontaneous recovery of 25% and a treatment effect of trimethoprim/sulfamethoxazole of 50% (based on a retrospective study of children treated with trimethoprim/sulfamethoxazole for COM at our hospital) and using an \( \alpha = 0.05 \) and a power of 0.80, we calculated that each group should consist of 50 children. Rate differences (RDs) with 95% confidence intervals (CIs) were calculated at the 3 control visits, to compare the 2 groups for otomicroscopic results, use of medication other than the study medication for ear disease, and adverse effects of the study medication. To detect possible effect modification, subgroup analyses were performed according to age (≤3 years or >3 years) and duration of otorrhea before study entry (≤6 months or >6 months), as prespecified in the trial protocol. Health-related quality-of-life instrument scores were transformed linearly onto scales of 0 to 100. The differences between the scores at the follow-up visits and at study entry were calculated and presented for each domain. Differences in domain scores between the groups at follow-up times of 0 weeks, 6 weeks, 12 weeks, and 1 year were tested with the Mann-Whitney \( U \) test, because these scores were not normally distributed. Box and whisker plots were used to compare the pure-tone hearing levels (air conduction at 500, 1000, 2000, and 4000 Hz) between the 2 groups. Percentage differences with 95% CIs were calculated for the bacteriologic findings. All analyses were performed on an intention-to-treat basis.

RESULTS

Study Group

Between February 2003 and November 2005, 101 children were enrolled; 50 were allocated to the trimethoprim/sulfamethoxazole group and 51 to the placebo group. The flow of the participants through the trial is presented in Fig 1. At baseline, clinical characteristics did not differ significantly between the 2 groups (Table 1). The median age of the study participants was 50 months (interquartile range: 55 months). The compliance rates for both the trimethoprim/sulfamethoxazole group and the placebo group were good (ie, >90% of the prescribed study medication was used).

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Characteristics of Patients With COM in the Trimethoprim/Sulfamethoxazole Group and the Placebo Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Characteristics</td>
</tr>
<tr>
<td></td>
<td>Trimethoprim/ Sulfamethoxazole ( (n = 50) )</td>
</tr>
<tr>
<td></td>
<td>Placebo ( (n = 51) )</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>28 (56)</td>
</tr>
<tr>
<td>Age, median (range), mo</td>
<td>48 (12–144)</td>
</tr>
<tr>
<td>Duration of otorrhea before study entry, median (range), mo</td>
<td>8 (3–113)</td>
</tr>
<tr>
<td>Previous treatment, n (%)</td>
<td>50 (100)</td>
</tr>
<tr>
<td>Ototopical drops</td>
<td>48 (96)</td>
</tr>
<tr>
<td>Systemic antibiotic therapy</td>
<td>30 (60)</td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
</tr>
<tr>
<td>Adenoidectomy and/or tonsillectomy</td>
<td>45 (90)</td>
</tr>
<tr>
<td>Tympanostomy tubes</td>
<td>5 (10)</td>
</tr>
<tr>
<td>Tympanoplasty and/or mastoidectomy</td>
<td>2 (0–6)</td>
</tr>
<tr>
<td>No. of siblings, median (range)</td>
<td>29 (58)</td>
</tr>
<tr>
<td>Family history of otitis media (parents or siblings), n (%)</td>
<td>6 (0–8)</td>
</tr>
<tr>
<td>Use of systemic antibiotic therapy during past 2 wk, n (%)</td>
<td>45 (90)</td>
</tr>
<tr>
<td>Parental smoking, n (%)</td>
<td>16 (32)</td>
</tr>
<tr>
<td>Day care or school in year before study entry, n (%)</td>
<td>3 (6)</td>
</tr>
<tr>
<td>Unilateral/bilateral COM, n (%)</td>
<td>25 (50)/25 (50)</td>
</tr>
<tr>
<td>Tympanostomy tubes, n (%)</td>
<td>32 (64)</td>
</tr>
<tr>
<td>Tympanic membrane perforation, n (%)</td>
<td>24 (48)</td>
</tr>
</tbody>
</table>

Eleven children (6 in the trimethoprim/sulfamethoxazole group and 5 in the placebo group) had a tympanostomy tube in one ear and a tympanic membrane perforation in the other.

Outcomes

At the 6-week follow-up visit, otorrhea was present for 28% of children in the trimethoprim/sulfamethoxazole group and 53% in the placebo group (RD: −25% to −6%; number needed to treat [NNT]: 4 children) (Table 2). At 12 weeks, the RD was still −15% (95% CI: −34% to 4%; NNT: 7 children) in favor of the trimethoprim/sulfamethoxazole group. At 1 year, there was no difference between the 2 groups (RD: 5%; 95% CI: −12% to 22%). At follow-up times of 6 weeks, 12 weeks, and 1 year, more children in the trimethoprim/sulfamethoxazole group had bilateral intact tympanic membranes and aerated middle ears than did those in the placebo group.

Otomicroscopic results were also analyzed according to age and duration of otorrhea before study entry. At
the 6-week follow-up visit, otorrhea was present in 8 children (29%) with >6 months of otorrhea before study entry who were treated with trimethoprim/sulfamethoxazole and 14 (67%) who were treated with placebo (RD: −38%; 95% CI: −64% to −12%; NNT: 3 children). For the children with 3 to 6 months of otorrhea before study entry, these values were 5 (26%) versus 13 (43%) (RD: −17%; 95% CI: −44% to 10%; NNT: 6 children). At the 12-week follow-up visit, these numbers were 11 (39%) versus 15 (71%) (RD: −32%; 95% CI: −59% to −5%; NNT: 3 children) and 4 (21%) versus 8 (29%) (RD: −8%; 95% CI: −33% to 17%; NNT: 13 children), respectively. At the follow-up time of 1 year, the treatment effect was no longer affected by the duration of otorrhea before study entry. Age did not influence the effectiveness of trimethoprim/sulfamethoxazole.

During the first 6 weeks, 38 (83%) of the children in the trimethoprim/sulfamethoxazole group and 39 (77%) of the children in the placebo group used antibiotic eardrops (RD: 6%; 95% CI: −10% to 22%). After the first 6 weeks, antibiotic eardrops were used slightly more often in the placebo group than in the trimethoprim/sulfamethoxazole group: 21 (55%) in the trimethoprim/sulfamethoxazole group and 26 (67%) in the placebo group (RD: −12%; 95% CI: −34% to 10%) used eardrops between follow-up times of 6 weeks and 12 weeks. Between follow-up times of 12 weeks and 1 year, these figures were 78% and 82% (RD: −4%; 95% CI: −22% to 14%), respectively. Systemic antibiotics other than the study medication (mostly amoxicillin) were used by 4 children (11%) in the trimethoprim/sulfamethoxazole group and 7 children (18%) in the placebo group between inclusion and the follow-up time of 12 weeks (RD: −7%; 95% CI: −23% to 9%). Between 12 weeks and 1 year, these figures were 23 (62%) and 18 (47%), respectively (RD: 15%; 95% CI: −7% to 37%), and trimethoprim/sulfamethoxazole and azithromycin were used most frequently. Ear/nose/throat surgery (tympanostomy tube insertion or removal, 6 children; adenotonsillectomy, 7; tympanomastoidectomy and/or tympanoplasty, 5) was performed for similar numbers of children in the 2 groups between follow-up times of 12 weeks and 1 year, that is, 13 (30%) in the trimethoprim/sulfamethoxazole group and 11 (24%) in the placebo group (RD: 6%; 95% CI: −12% to 24%).

Figure 2 shows box and whisker plots of the hearing levels for children >3 years of age. Pure-tone air conduction levels at 500, 1000, 2000, and 4000 Hz could be determined for 20 children in the trimethoprim/sulfamethoxazole group and 18 children in the placebo group. Although hearing levels generally improved, no differences between the groups were found.

During the study, the health-related quality-of-life scores improved substantially in both the trimethoprim/sulfamethoxazole and placebo groups (data not shown). Mean scores for the trimethoprim/sulfamethoxazole and placebo groups for the 6-item otitis media questionnaire, Child Health Questionnaire, and visual analog scale were the same at all visits.

Table 3 shows that, at follow-up times of 6 weeks, 12 weeks, and 1 year, there were no differences in the proportions of culture-positive otorrhea samples for the trimethoprim/sulfamethoxazole group and the placebo group (6 weeks: RD: 0%; 95% CI: −14% to 14%; 12 weeks: RD: 3%; 95% CI: −20% to 26%; 1 year: RD: −8%; 95% CI: −23% to 7%). At follow-up times of 6 and 12 weeks, Pseudomonas aeruginosa was the most frequently isolated microorganism in both groups and, in contrast to most other organisms, was found more frequently in the otorrhea samples of the trimethoprim/sulfamethoxazole group than in those of the placebo group, that is, 9 (56%) vs 12 (38%) at 6 weeks (RD: 18%; 95% CI: −12% to 48%) and 7 (50%) vs 9 (31%) at 12 weeks (RD: 19%; 95% CI: −12% to 50%). At the follow-up time of 1 year, no differences were found between the groups.

**Adverse Effects**

During the first 6 weeks, vomiting or diarrhea (potential adverse effects of the study medication) were reported for 9% of the trimethoprim/sulfamethoxazole group and 2% of the placebo group (RD: 7%; 95% CI: −2% to 16%; number needed to harm: 14 children). Between 6
and 12 weeks, no differences between the groups were found. One child in the trimethoprim/sulfamethoxazole group developed a skin rash; she was treated with cetirizine and administration of the study medication was discontinued, after which the rash disappeared. Treatment of COM was continued with azithromycin. Complete blood count and hepatic and renal function tests at inclusion and at follow-up times of 6 weeks and 12 weeks did not show any abnormalities in either group.

Two children developed mastoiditis during the first 12 weeks of follow-up monitoring, and their randomization codes were unblinded immediately. One child was allocated to trimethoprim/sulfamethoxazole; he was treated with a mastoidectomy and amoxicillin/clavulanic acid, administered intravenously for 7 days and orally for 14 days. The other child was allocated to placebo; he was treated with intravenously administered amoxicillin/clavulanic acid for 7 days, followed by 6 weeks of trimethoprim/sulfamethoxazole. Both children recovered well with this therapy.

**DISCUSSION**

This is the first placebo-controlled, randomized trial of systemic antibiotic treatment for patients with COM. It showed that a 6- to 12-week high-dose course of trimethoprim/sulfamethoxazole in addition to a short course of steroid and antibiotic eardrops had a cure rate of 68% at the follow-up time of 12 weeks and was clinically more effective than placebo for children with COM who had experienced failure of conventional management with topical medications or short-term systemic antibiotic therapy. This treatment effect was most pronounced during the first 6 weeks. Children with a history of otorrhea for >6 months benefited more from trimethoprim/sulfamethoxazole therapy than did those with a shorter history of otorrhea. Age did not influence the effectiveness of trimethoprim/sulfamethoxazole.

Pure-tone hearing levels and health-related quality of life improved during the study but did not differ between the trimethoprim/sulfamethoxazole group and the placebo group.

The effect of treatment with trimethoprim/sulfamethoxazole for COM was studied in one previous trial, in which a 2-week course of trimethoprim/sulfamethoxazole was compared with a course of antibiotics based on culture results. After a maximal follow-up period of only 14 days, otorrhea resolved for 75% of the trimethoprim/sulfamethoxazole group and 85% of the patients treated with culture-directed antibiotics. Other studies using various systemic antibiotics for COM found similar success rates of ~70%. Inclusion and outcome criteria, route of drug administration, and follow-up periods in those studies varied considerably, and no study was placebo controlled.

Our results need to be interpreted in light of several limitations. First, the children in our study had persistent symptoms of COM despite conventional management with topical medications, short-term systemic antibiotic therapy, and/or previous ear surgery. Because the majority of children with otorrhea seen by generalist physicians respond well to conventional management, our results should be applied to patients with similarly complicated COM.

Second, because the Netherlands is known for its...
⫺8 (⫺23 to 7)
⫺1 (⫺23 to 21)
8 (⫺7 to 23)
6 (⫺20 to 32)
6 (⫺20 to 32)
⫺9 (⫺49 to 31)
0 (0 to 0)
⫺28 (⫺63 to 7)
11 (100)
1 (9)
0 (0)
1 (9)
1 (9)
6 (55)
0 (0)
9 (82)
18
12 (92)
1 (8)
1 (8)
2 (15)
2 (15)
6 (46)
0 (0)
7 (54)
19
3 (⫺20 to 26)
⫺14 (⫺27 to ⫺1)
⫺3 (⫺20 to 14)
⫺24 (⫺45 to ⫺3)
⫺14 (⫺34 to 6)
19 (⫺12 to 50)
⫺10 (⫺21 to 1)
⫺5 (⫺34 to 24)
24 (83)
4 (14)
3 (10)
9 (31)
6 (21)
9 (31)
3 (10)
10 (34)
44
12 (86)
0 (0)
1 (7)
1 (7)
1 (7)
7 (50)
0 (0)
4 (29)
14
0 (⫺14 to 14)
⫺16 (⫺29 to ⫺3)
⫺6 (⫺14 to 2)
⫺29 (⫺45 to ⫺13)
⫺6 (⫺27 to 15)
18 (⫺12 to 48)
⫺3 (⫺9 to 3)
25 (⫺4 to 54)
30 (94)
5 (16)
2 (6)
6 (29)
6 (19)
12 (38)
1 (3)
12 (38)
44
15 (94)
0 (0)
0 (0)
0 (0)
2 (13)
9 (56)
0 (0)
10 (63)
21
TMP/SMX indicates trimethoprim/sulfamethoxazole.

Placebo
(n ⫽ 32)
TMP/SMX
(n ⫽ 16)
Placebo
(n ⫽ 70)
TMP/SMX
(n ⫽ 74)

59 (84)
7 (10)
1 (1)
23 (33)
14 (20)
22 (31)
6 (9)
23 (33)
95

Placebo
(n ⫽ 29)
TMP/SMX
(n ⫽ 14)

No. (%)

RD
(95% CI), %
No. (%)
No. (%)

68 (92)
10 (14)
4 (5)
13 (18)
15 (20)
32 (43)
1 (1)
53 (72)
128
Positive culture
Streptococcus pneumoniae
Hemolytic streptococci, group A
Haemophilus inﬂuenzae
Staphylococcus aureus
Pseudomonas aeruginosa
Moraxella catarrhalis
Other organisms
Total No. of species

TMP/SMX
(n ⫽ 13)

Placebo
(n ⫽ 11)

RD
(95% CI), %
RD
(95% CI), %

No. (%)

1y
12 wk
6 wk
Inclusion
Bacterial Species

TABLE 3 Bacteriologic Findings for the Otorrhea Samples of Children at Inclusion and During Follow-Up Monitoring (per Sample Taken)

restrictive policy regarding systemic antibiotic treatment
for otitis media, it is possible that before study entry the
participants had received fewer courses of systemic antibiotics and more courses of topical antibiotics than
would be expected in other countries. A meta-analysis
by Macfadyen et al,16 however, showed that topical antibiotics, such as those used by our participants before
study entry, were more effective than short courses of
systemic antibiotic therapy in resolving otorrhea. Therefore, we think that our results can be extrapolated to
countries where short-term systemic antibiotic therapy
is used more frequently for the management of COM.
Third, children in our study all received suction cleaning and topical treatment, in addition to the study medication, when otorrhea was present. Antibiotic eardrops
were used slightly more frequently in the placebo group
than in the trimethoprim/sulfamethoxazole group between follow-up times of 6 and 12 weeks. This might
have influenced the high cure rate in the placebo group
at 12 weeks, which might have resulted in an underestimation of the real treatment effect of trimethoprim/
sulfamethoxazole. Other important factors for the small
treatment effect with the longer course might have been
the natural course of COM and regression to the mean.
Fourth, at the follow-up time of 12 weeks, the parents
and the local otolaryngologist or pediatrician were informed about the assigned treatment, and doctors were
free to manage additional symptoms of COM in both
groups either according to our advice (with a 6-week
course of antibiotics) or according to their own practice.
Our follow-up data revealed that 10 children in the
placebo group and 7 children in the trimethoprim/sulfamethoxazole group indeed received a prolonged
course of antibiotics after the follow-up period of 12
weeks. This might have added to the dilution of the
effect after 12 weeks.
Fifth, the choice of trimethoprim/sulfamethoxazole
for COM could be questioned, because P aeruginosa,
which is the most common organism in COM and was
present in 54 (38%) of otorrhea samples at inclusion, is
known to be unsusceptible to trimethoprim/sulfamethoxazole. This is reflected by our culture results; during
treatment with trimethoprim/sulfamethoxazole, the
proportion of otorrhea samples positive for P aeruginosa
did not change, whereas that of bacteria that are not
intrinsically resistant to trimethoprim/sulfamethoxazole,
such as H influenzae and S pneumoniae, decreased. Because trimethoprim/sulfamethoxazole was effective during the first 6 to 12 weeks of our study, P aeruginosa
seems to be a secondary microorganism in COM, rather
than the causative microorganism.
CONCLUSIONS
A 6- to 12-week course of high-dose, oral, trimethoprim/sulfamethoxazole therapy is beneficial for
children suffering from COM. The treatment effect is
PEDIATRICS Volume 119, Number 5, May 2007

903


ACKNOWLEDGMENTS
This study was funded by grant 01-235 from the Health Care Insurance Council, Netherlands.

We thank A. McArthur-Noom for the microbiologic determinations and P. M. van Eden for conscientious secretarial support.

REFERENCES

RANDOMIZED CLINICAL TRIALS
Pediatrics requires investigators to register their clinical trials in a public trials registry. The members of the International Committee of Medical Journal Editors (ICMJE) will consider most clinical trials for publication only if they have been registered (see N Engl J Med. 2004;351:1250–1251). Current information on requirements and appropriate registries is available at www.icmje.org/faq.pdf.
Aminoglycoside-Based Triple-Antibiotic Therapy Versus Monotherapy for Children With Ruptured Appendicitis

Adam B. Goldin, MD, MPH, Robert S. Sawin, MD, MS, Michelle M. Garrison, PhD, Danielle M. Zerr, MD, MPH, Dimitri A. Christakis, MD, MPH

Departments of *Pediatric General and Thoracic Surgery, †Pediatric Infectious Diseases, and ‡Pediatrics, Children’s Hospital and Regional Medical Center, Seattle, Washington; §Child Health Institute, University of Washington Medical Center, Seattle, Washington

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

ABSTRACT

OBJECTIVE. We conducted a retrospective cohort study to compare the use of triple therapy versus monotherapy for children and adolescents with perforated appendicitis and to determine whether there has been a transition to monotherapy within the freestanding children’s hospitals that contribute to the Pediatric Health Information System database.

METHODS. We used the Pediatric Health Information System database, which includes billing and discharge data for 32 children’s hospitals in the United States, to examine the trend in antibiotic usage and whether the postappendectomy antibiotic regimen was associated with differences in complication-related readmissions, length of stay, or charges in a population of children and adolescents with ruptured appendicitis and discharge dates between March 1, 1999, and September 30, 2004. Pairwise regression analyses were performed to compare the most common monotherapy regimens with the triple therapy.

RESULTS. A total of 8545 patients met the inclusion criteria, of whom 58%, over the entire study period, received the aminoglycoside-based triple antibiotic therapy on postoperative day 1. There was, however, a notable transition over this 6-year period, from 69% to 52% of surgeons using aminoglycoside-based combination therapy. There were no significant differences in the odds of readmission at 30 days except for the group receiving ceftriaxone, which was associated with significantly decreased odds. The subgroup receiving piperacillin/tazobactam monotherapy demonstrated significantly decreased length of stay (−0.90 days) and total hospital charges, and the group receiving cefoxitin demonstrated significantly decreased length of stay (−1.89 days), as well as decreased pharmacy and total hospital charges.

CONCLUSIONS. Single-agent antibiotic therapy in the treatment of perforated appendicitis is being used with increasing frequency, is at least equal in efficacy to the traditional aminoglycoside-based combination therapy, and may offer improvements in terms of length of stay, pharmacy charges, and hospital charges.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2040
doi:10.1542/peds.2006-2040

Key Words
appendicitis, antibiotics

Abbreviations
PHIS—Pediatric Health Information System
OR—odds ratio
CI—confidence interval
ICD-9—International Classification of Diseases, Ninth Revision
LOS—length of stay
ABT—aminoglycoside-based combination therapy

Accepted for publication Jan 8, 2007

Address correspondence to Adam B. Goldin, MD, MPH, Department of Pediatric General and Thoracic Surgery, Children’s Hospital and Regional Medical Center, 4800 Sand Point Way NE, Seattle, WA 98105. E-mail: adam.goldin@seattlechildrens.org

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
Approximately 4 appendectomies per 1000 children are performed each year in the United States. Of these cases, 15% to 36% are in the presence of perforation.\textsuperscript{1-3} Although intravenous antibiotic therapy is a cornerstone of therapy for perforated appendicitis, the optimal regimen of antibiotics for children remains controversial. Although many physicians would argue that “standard therapy” consists of an aminoglycoside, a β-lactam, and an antibiotic with anaerobe coverage (ie, ampicillin, gentamicin, or metronidazole),\textsuperscript{4-16} increasing evidence suggests that single-agent antibiotic therapy provides equivalent results, compared with multiagent regimens.\textsuperscript{17-32} In fact, a meta-analysis reviewed 64 randomized, controlled trials comparing β-lactam monotherapy with β-lactam/aminoglycoside combination therapy for patients with sepsis.\textsuperscript{33} The addition of an aminoglycoside increased the risk for adverse events while leaving fatality rates unchanged. This has led some to conclude that, for sepsis, adding an aminoglycoside to β-lactams should be discouraged.\textsuperscript{33,34} Although these studies are not specific to ruptured appendicitis in children, they do call into question the use of aminoglycoside-based combination therapy (ABT). Given the expanding body of evidence suggesting that single-agent, broad-spectrum, antibiotic therapy for children with perforated appendicitis is both safe and effective, we conducted a retrospective cohort study to compare the use of ABT with that of monotherapy for children and adolescents with perforated appendicitis and to determine whether there has been a transition to monotherapy within the freestanding children’s hospitals throughout the nation that contribute to the Pediatric Health Information System (PHIS) database. The study end points included hospital readmissions because of early and late complications of perforated appendicitis, as well as the length of stay (LOS) and hospital charges for the initial admission.

Patients
Our study included pediatric patients (through 18 years of age) with discharge dates between March 1, 1999, and September 30, 2004. We restricted the analysis to patients who had both an ICD-9 diagnosis code for ruptured appendicitis (540.0x or 540.1x) and an ICD-9 procedure code (47.01, 47.09, 47.2x, or 47.99) for appendectomy during the same hospital stay. Unique patient identifiers were used to ensure that each patient was included in the study only once.

Primary Outcomes and Predictors
The primary outcome measure in the analysis was complication-related readmissions. Because of the nature of the data, we were not able to assess complications that might have been diagnosed during the initial admission. Complications were determined by using discharge diagnosis codes, which are associated with the admission as a whole, rather than with any specific date within the stay. Therefore, it was not possible to ascertain whether such conditions arose before or after the appendectomy procedure; clearly, conditions that arose before the appendectomy and the antibiotic protocol could not legitimately be considered complications. Readmissions were categorized as early (postoperative day <31) or late (postoperative day 31–730) and were analyzed separately because of an a priori decision that the mechanisms involved would differ. Complication-related readmissions were defined as those with primary ICD-9 diagnoses of peritonitis, appendicitis or appendiceal abscess, surgical site infection, Clostridium difficile infection, postsurgical complications, intestinal obstruction or adhesion, or abdominal pain. The timing of early and late complications was defined on the basis of previously published intervals for early and late adhesive bowel obstruction.\textsuperscript{35,36} Secondary outcomes examined were postoperative LOS, pharmacy charges during the hospital stay, and total charges during the hospital stay. For both charge-related outcomes, we used measures of charges that were adjusted by using the Centers for Medicare and Medicaid Services wage/price index for the hospital’s location.

Patients were categorized into 1 of 5 groups on the basis of antibiotics administered on postoperative day 1 (Table 1). This day was chosen to ensure that preoperative antibiotics were not captured, thereby maintaining an intent-to-treat perspective, because the surgeons’ knowledge of perforated versus nonperforated appendicitis would be definite. The reference group in each analysis was composed of patients receiving ABT and was examined in pairwise comparisons with the 4 most common types of monotherapies used. Patients were not included in a category if on postoperative day 1 they received antibiotics in addition to those specified for the category. Only patients who met the criteria for 1 of the 5 antibiotic groups were included in the analysis. Be-

METHODS
Data Source
We used the PHIS database developed by the Child Health Corporation of America, which includes demographic, diagnostic, and charge data for freestanding, noncompeting, children’s hospitals. The PHIS includes both diagnoses and procedures coded by using the International Classification of Diseases, Ninth Revision (ICD-9) format. For a subset of participating hospitals, the database also has level II data, which use Clinical Transaction Classification codes to map hospital-specific charge codes at the patient level to categorical variables across all hospitals; because medication use data are available only for this subset, we restricted our analysis to the 32 hospitals that had level II data coded during the study period.
cause ampicillin/sulbactam does not have equivalent antimicrobial coverage, compared with the other antibiotics in the piperacillin/tazobactam category, we decided to perform a subanalysis to determine whether the results for that category differed if patients receiving ampicillin/sulbactam were removed.

**Statistical Analyses**

Descriptive characteristics of the population were compared between the groups receiving ABT versus any of the monotherapy regimens, by using χ² tests for dichotomous variables (gender, Medicaid status, laparoscopic procedure, abscess diagnosed, and procedure day) and t tests for continuous variables (age, LOS, and procedure day). In the multivariate analyses, we used logistic regression for the readmission outcomes and linear regression for LOS and charge outcomes. Although the distributions of both charges and LOS were highly skewed, our large sample size made linear regression appropriate. All regressions were performed pairwise, comparing each of the monotherapy groups in turn with the ABT group as reference.

We controlled for potential confounders by using a propensity score, which was developed separately for each pairwise comparison by using hierarchical, stepwise, logistic regression. Only covariates that likely would have been known to the providers at initial antibiotic assignment on postoperative day 1 were included in the models to develop the propensity scores. The covariates entered into the stepwise model included demographic variables (gender, age, and Medicaid status), clinical factors ascertainable at antibiotic assignment (chronic comorbid conditions, priority of admission score, abscess diagnosed, procedure performed laparoscopically, and day of hospital stay on which procedure was performed), and hospital stay variables (hospital and admission date). Predicted probability of assignment to the given monotherapy group was used as the propensity score covariate in the analytic regressions.

We also stratified the readmission regressions according to whether an abscess was diagnosed during the initial hospitalization for ruptured appendicitis, because we thought that the effect of the antibiotics could differ across the 2 populations. In addition, a covariate for follow-up time was included in the regressions examining readmissions as an outcome, because fewer months of follow-up data were available for index hospitalizations that occurred later in the study period. The follow-up time variable was modeled as the percentage of time examined for readmission (30 or 720 days, respectively) that could have been captured after each index hospitalization. Finally, all regression analyses were clustered with respect to hospital, to account for the decreased within-hospital variation, as opposed to that between hospitals.

To protect the integrity of the participating hospitals, all results are presented with hospitals deidentified. All analyses were conducted by using Stata 8.0 software (Stata, College Station, TX).

**RESULTS**

Of the 14,616 hospital admissions with a ruptured appendicitis diagnosis and appendectomy procedure coded in the PHIS database, 8,545 were unique patients who met all inclusion criteria for the analysis (Fig 1). Of those, 58% received ABT on postoperative day 1. The most common monotherapy was second-generation cephalosporins such as cefotixin. Regimen selection did vary over time, with 69% of study patients admitted in 1999 receiving the triple therapy, compared with 52% in 2004. Significant increases over this time period in the use of each monotherapy, except for the cefotixin group, were observed.

Although the only significant difference between groups in the demographic variables was a slightly lower mean age in the triple-therapy group (Table 2), significant differences between groups were observed across every appendectomy-related variable. Patients who received triple therapy were significantly less likely to have had a laparoscopic procedure (27% vs 57%) and were significantly more likely to have had an abscess diagnosed (30% vs 25%) than were those who received a monotherapy regimen. A somewhat longer mean LOS was observed for the triple-therapy group, compared with the monotherapy groups (6.83 vs 5.77 days).

In the pairwise regressions examining readmissions (Table 3), the ceftriaxone group was associated with reduced odds for 30-day readmission, both in the complete sample (odds ratio [OR]: 0.56; 95% confidence interval [CI]: 0.34–0.91) and in the sample without an abscess diagnosis (OR: 0.27; 95% CI: 0.12–0.60), compared with the triple-therapy group. Similarly, the piperacillin/tazobactam group was associated with decreased odds of 2-year readmission in the sample without an abscess diagnosis (OR: 0.36; 95% CI: 0.19–0.67). However, both the metronem group and the

**TABLE 1 Antibiotic Categories**

<table>
<thead>
<tr>
<th>ABT regimen</th>
<th>Antibiotic monotherapy regimens</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABT</td>
<td>Piperacillin/tazobactam, etc</td>
</tr>
<tr>
<td></td>
<td>Meropenem, etc</td>
</tr>
<tr>
<td></td>
<td>Ceftriaxone, etc</td>
</tr>
</tbody>
</table>

a All analyses examining the piperacillin/tazobactam category were also performed after exclusion of ampicillin/sulbactam.
Ceftriaxone group seemed to be associated with increased odds for 2-year readmission, with significant findings in the complete sample for meropenem (OR: 3.74; 95% CI: 1.08–12.94). For the ceftriaxone group, the findings were statistically significant in both the complete sample (OR: 2.69; 95% CI: 1.17–6.21) and the sample with an abscess diagnosis (OR: 2.43; 95% CI: 1.24–4.76). No significant changes in the findings were observed when regressions were repeated after removal of the patients receiving ampicillin/sulbactam from the piperacillin/tazobactam group.

In the regressions examining antibiotic regimen and postoperative LOS (Table 4), significant decreases were seen in the cefoxitin (−1.89 days; 95% CI: −2.75 to −1.03 days) and piperacillin/tazobactam (−0.90 days; 95% CI: −1.69 to −0.11 days) groups, compared with the triple-therapy group. Cefoxitin was also associated with significantly decreased pharmacy charges (−2187 dollars; 95% CI: −3110 to −1264 dollars), and both cefoxitin and piperacillin/tazobactam were associated with decreased total charges, compared with the triple-therapy group.

**DISCUSSION**

Appendicitis is the most common indication for urgent abdominal surgery in the pediatric population, and ruptured appendicitis affects a large proportion of those patients. Minimizations of morbidity, costs, hospital LOS, and readmissions remain primary objectives of surgical management. Postappendectomy infection relates to all of these measurable outcomes, and the choice of care.
antibiotic regimens has a major effect on each of these parameters.

The optimal antibiotic regimen for the treatment of ruptured appendicitis, however, has yet to be determined. As outlined above, many recent articles have suggested that monotherapy may be preferable to the traditional multiagent regimen, and the evidence presented herein demonstrates that monotherapy is being used with increasing frequency in the treatment of perforated appendicitis. In addition, our findings support many of the aforementioned studies that do not demonstrate a difference in the effectiveness of ABT, compared with monotherapy, for either early or late readmissions. The differences in our study that were significant were not consistent across early and late readmissions. The findings from our study demonstrated that, although there was little difference in antibiotic regimens in terms of odds of readmission, there were significant differences in terms of postoperative LOS, pharmacy charges, and total hospitalization charges over a 5-year period among 32 freestanding pediatric hospitals throughout the nation. These findings are important in that they support the use of single-agent antibiotic therapy as a safe and more cost-effective alternative to ABT for the treatment of perforated appendicitis.

Our findings demonstrated a transition toward a preference for monotherapy in the 5 years during which we evaluated patients in the PHIS. Whereas the frequency of ABT was 69% during the first year of our study, 52% of patients received monotherapy during the last year. Articles evaluating the treatment of sepsis in adults have demonstrated improved outcomes with monotherapy, compared with ABT,33,34 and others have demonstrated improved outcomes in randomized trials of treatment of perforated appendicitis in children.20–24,26,28–32 Despite existing evidence, however, tremendous variability in practice patterns, including the use of antibiotics, among pediatric surgeons still endures.3 Almost 60% of surgeons base their clinical practice in the management of perforated appendicitis on their individual preferences.38

There are several limitations to our analysis. First, the PHIS is an administrative discharge database that requires data entry by nonmedical staff members; therefore, there is potential for misclassification. The principal diagnoses used, as well as the medications dispensed, however, are likely to have been coded accurately. Moreover, any random misclassification would bias our findings toward the null hypothesis. There is no a priori reason to suspect systematic bias in the coding of complications and antibiotic regimens. Second, the observational nature of this study precludes drawing causal inferences. For example, we do not know whether the

### TABLE 2 Description of Study Sample

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Total</th>
<th>Without Abscess</th>
<th>With Abscess</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>95% CI</td>
<td>P</td>
<td>OR</td>
</tr>
<tr>
<td>Piperacillin/tazobactam</td>
<td>1.17</td>
<td>0.61–2.22</td>
<td>.63</td>
</tr>
<tr>
<td>Meropenem</td>
<td>0.82</td>
<td>0.53–1.26</td>
<td>.36</td>
</tr>
<tr>
<td>Cefoxitin</td>
<td>1.22</td>
<td>0.79–1.89</td>
<td>.37</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>0.56*</td>
<td>0.34–0.91</td>
<td>.02</td>
</tr>
<tr>
<td>Piperacillin/tazobactam</td>
<td>1.01</td>
<td>0.35–2.93</td>
<td>.99</td>
</tr>
<tr>
<td>Meropenem</td>
<td>3.74*</td>
<td>1.08–12.94</td>
<td>.04</td>
</tr>
<tr>
<td>Cefoxitin</td>
<td>0.69</td>
<td>0.42–1.13</td>
<td>.14</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>2.69*</td>
<td>1.17–6.21</td>
<td>.02</td>
</tr>
</tbody>
</table>

Monotherapy was compared with triple therapy in pairwise regressions. All models were adjusted for propensity score and follow-up time.

* ORs were statistically significant at the P < .05 level.
TABLE 4  Monotherapy Versus Triple Therapy and Hospital Stay Characteristics

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Difference in LOS or Cost</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postoperative LOS, d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piperacillin/tazobactam</td>
<td>−0.90*</td>
<td>−1.69 to −0.11</td>
<td>.03</td>
</tr>
<tr>
<td>Meropenem</td>
<td>−1.09</td>
<td>−3.37 to 1.19</td>
<td>.33</td>
</tr>
<tr>
<td>Cefoxitin</td>
<td>−1.89*</td>
<td>−2.75 to −1.03</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>−1.15</td>
<td>−2.42 to 0.13</td>
<td>.08</td>
</tr>
<tr>
<td>Pharmacy charges, $</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piperacillin/tazobactam</td>
<td>−1147</td>
<td>−2328 to 33</td>
<td>.06</td>
</tr>
<tr>
<td>Meropenem</td>
<td>1563</td>
<td>−1498 to 4624</td>
<td>.31</td>
</tr>
<tr>
<td>Cefoxitin</td>
<td>−2187*</td>
<td>−3110 to −1264</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>−353</td>
<td>−2796 to 2090</td>
<td>.77</td>
</tr>
<tr>
<td>Total charges, $</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piperacillin/tazobactam</td>
<td>−4829*</td>
<td>−8431 to −1227</td>
<td>.01</td>
</tr>
<tr>
<td>Meropenem</td>
<td>−1389</td>
<td>−12 732 to 9954</td>
<td>.80</td>
</tr>
<tr>
<td>Cefoxitin</td>
<td>−6728*</td>
<td>−10 120 to −3336</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>−5622</td>
<td>−12 473 to 1229</td>
<td>.10</td>
</tr>
</tbody>
</table>

* Effect estimates were statistically significant at the P < .05 level.

LOS for patients who received monotherapy was shorter because of the choice of antibiotics or because surgeons who tend to choose monotherapy also tend to be more aggressive about discharging patients from the hospital. The purpose of an observational study is to identify associations that will allow us to develop well-designed, hypothesis-driven studies. Third, our data set includes only patients treated in the 32 hospitals in the PHIS database. Although this represents a large diverse sample of patients and institutions, the extent to which these findings can be generalized to other community-based institutions is unknown. Fourth, we were not able to address the trend toward the use of interval appendectomy in this study. Because our inclusion criteria required codes for both diagnosis of appendicitis and the procedure of an appendectomy, we did not include patients who were admitted with perforated appendicitis and treated with antibiotics or percutaneous drainage and antibiotics. In addition, the second hospitalization with a diagnosis of appendicitis was excluded; therefore, those patients would not have been included on their return for their operations. Fifth, we did not try to identify the frequency of central line placement and home intravenous antibiotic treatment. This is a common mode of treatment that influences the evaluation of outcomes such as hospital costs and LOS. Sixth, almost 42% of the patients in the database were excluded because they did not meet our inclusion criteria, because of age, because it was a repeat admission with appendicitis, because the patient received care at an institution that did not contribute pharmacy data to the PHIS, or because the patient received an antibiotic that did not fall into one of our predefined regimens. The majority of those patients were excluded on the basis of available data; therefore, we assumed that this would not introduce a significant bias into the analysis. The patients who were excluded because they received additional antibiotics on postoperative day 1 (17% of the population), however, remain more difficult to describe. Those patients were excluded primarily because they could not be assigned accurately to one of our predetermined groups. The patients received unusual combinations of drugs, and there were not enough patients with each of the combinations for analysis as a group. In our study, the least commonly used monotherapy regimen was ceftriaxone; this population represented only 3% of the total population. We elected to concentrate our analysis on the most commonly used regimens, rather than adding more data for <1% of the population. Despite these limitations, our data suggest that single-agent therapy for ruptured appendicitis in children is accepted practice among almost one half of surgeons practicing at children’s hospitals contributing to the PHIS database and may be optimal with respect to morbidity, costs, and LOS.

REFERENCES


Determinants of Outcomes After Head Cooling for Neonatal Encephalopathy

John S. Wyatt, MBBSa, Peter D. Gluckman, FRScb, Ping Y. Liu, PhDc, Denis Azzopardi, MDd, Roberta Ballard, MD*, A. David Edwards, FMedSci*, Donna M. Ferriero, MDe, Richard A. Polin, MD*, Charlene M. Robertson, MD, Marianne Thoresen, MD, PhD, Andrew Whitelaw, MD, Alistair J. Gunn, MBChB, PhD*, for the CoolCap Study Group

aDepartment of Paediatrics and Child Health, University College London, London, United Kingdom; bThe Liggins Institute and Department of Physiology, University of Auckland, Auckland, New Zealand; cDrug and Device Development Company, Redmond, Washington; dDivision of Clinical Sciences, Imperial College, London, United Kingdom; eDepartment of Pediatrics, Children’s Hospital of Philadelphia, University of Pennsylvania, Philadelphia, Pennsylvania; Departments of Neurology and Pediatrics, University of California, San Francisco, San Francisco, California; fCollege of Physicians and Surgeons, Columbia University, New York, New York; Department of Pediatrics, University of Alberta, Edmonton, Canada; gDepartment of Clinical Science, St Michael’s Hospital, and hDepartment of Clinical Science, Southmead Hospital, University of Bristol, Bristol, United Kingdom

Financial Disclosure: Olympic Medical provided financial grants to the University of Auckland (Dr. Gluckman) and University College London (Dr. Wyatt) to cover the costs of administering the CoolCap trial. Each participating trial site (not the individual site investigators) received fixed partial reimbursement for each infant enrolled, covering the additional costs of performing the trial, including clinical time, laboratory tests, and neurodevelopmental assessments, and local trial administration. Dr. Azzopardi authored a manual of aEEG interpretation that was distributed by Olympic Medical for use in the trial. Olympic Medical loaned equipment to Drs. Gunn, Thoresen, Whitelaw, and Wyatt for pilot studies preceding the trial. The University of Auckland has applied for a related patent that names Dr. Gunn; however, Drs. Gunn and Gluckman have no financial interest.

ABSTRACT

OBJECTIVE. The goal of this study was to evaluate the role of factors that may determine the efficacy of treatment with delayed head cooling and mild systemic hypothermia for neonatal encephalopathy.

METHODS. A total of 218 term infants with moderate to severe neonatal encephalopathy plus abnormal amplitude-integrated electroencephalographic recordings, assigned randomly to head cooling for 72 hours, starting within 6 hours after birth (with the rectal temperature maintained at 34.5 ± 0.5°C), or conventional care, were studied. Death or severe disability at 18 months of age was assessed in a multicenter, randomized, controlled study (the CoolCap trial).

RESULTS. Treatment, lower encephalopathy grade, lower birth weight, greater amplitude-integrated electroencephalographic amplitude, absence of seizures, and higher Apgar score, but not gender or gestational age, were associated significantly with better outcomes. In a multivariate analysis, each of the individually predictive factors except for Apgar score remained predictive. There was a significant interaction between treatment and birth weight, categorized as ≥25th or <25th percentile for term, such that larger infants showed a lower frequency of favorable outcomes in the control group but greater improvement with cooling. For larger infants, the number needed to treat was 3.8. Pyrexia (≥38°C) in control infants was associated with adverse outcomes. Although there was a small correlation with birth weight, the adverse effect of greater birth weight in control infants remained significant after adjustment for pyrexia and severity of encephalopathy.

CONCLUSIONS. Outcomes after hypothermic treatment were strongly influenced by the severity of neonatal encephalopathy. The protective effect of hypothermia was greater in larger infants.
Neonatal encephalopathy associated with exposure to hypoxia/ischemia continues to be an important cause of acute neurologic injury, occurring in ~2 to 3 cases per 1000 term live births in developed countries, with a higher incidence in less developed countries. Recently, we reported in a multicenter, randomized trial (the CoolCap trial) that head cooling with mild systemic hypothermia was associated with an improved rate of survival, without severe disability at 17 to 22 months of age, for all infants except those defined a priori as having the most severe changes on the baseline amplitude-integrated electroencephalographic (aEEG) recordings. There was no increase in rates of significant adverse events for cooled infants. These findings are consistent with those for whole-body cooling. A recent meta-analysis of those studies confirmed a significant overall effect of postpartum hypothermia but noted that there was considerable potential for additional improvement, and it is not certain that the 2 approaches are comparable in either their long-term or short-term outcomes.

It is of great importance for future clinical and experimental studies to determine the clinical factors that might have influenced the effectiveness of treatment. We reported previously, for example, that baseline aEEG recordings could stratify patients as those who might respond and those for whom therapy would have no effect. The Apgar scores during resuscitation and the clinical severity of encephalopathy after birth are also predictive of adverse neurodevelopmental outcomes, but it is not known whether they modulate responses to treatment. Similarly, growth-restricted infants and boys have higher rates of encephalopathy, but it is not known whether these factors affect the severity of outcomes or influence therapeutic efficacy. It is possible, for example, that being born small is associated with perinatal insults that might compromise the brain before the peripartum period. In the CoolCap study, clinical grading of encephalopathy and aEEG recordings were obtained before randomization, to improve the specificity of case selection and to provide objective pretreatment data on the severity of brain injury. We now present the results of an exploratory analysis to examine a range of possible clinical factors that might influence outcomes after selective head cooling for treatment of neonatal encephalopathy, including the encephalopathy grade, 5-minute Apgar score, gender, birth weight, aEEG findings, and pyrexia in control infants during the 76-hour interval after randomization, to generate hypotheses for future studies.

Methods

Study Design

This study was performed in 25 perinatal centers with a trial design registered with the US Food and Drug Administration under the investigational device exemption/premarket approval program. The institutional review board of each center approved the protocol, and written informed consent was obtained from parents before randomization. From July 1999 to January 2002, infants at ≥36 weeks of gestation with acute encephalopathy were recruited by using a stepwise protocol including clinical evidence of exposure to perinatal hypoxia/ischemia, an abnormal neurologic examination, and an abnormal aEEG recording. The entry criteria, method of randomization, exclusion criteria, defined adverse events, data collection, independent data safety and monitoring board, general management, primary outcome of death or severe disability at 18 months of age, safety outcomes, and interaction between severity of aEEG changes and response to hypothermia were reported elsewhere.

Study Entry Criteria

The required clinical criteria were an Apgar score of ≤5 at 10 minutes after birth, continued need for resuscitation (including endotracheal or mask ventilation) at 10 minutes after birth, or severe acidosis (defined as either pH of <7.00 or base deficit of ≥16 mmol/L in an umbilical cord blood sample or an arterial or venous sample obtained within 60 minutes after birth). The infants were then assessed, by a certified examiner, for evidence of moderate (grade 2) or severe (grade 3) encephalopathy according to criteria modified from those described by Sarnat and Sarnat, including lethargy, stupor, or coma with ≥1 of hypotonia, abnormal reflexes including oculomotor or pupillary abnormalities, absent or weak suck, or clinical evidence of seizures.

Infants who satisfied the aforementioned criteria then underwent ≥20 minutes of aEEG recording (Lectromed, Letchworth, United Kingdom) by investigators. The aEEG recording could be performed at any time after 1 hour of age, except within 30 minutes after intravenous administration of anticonvulsant therapy, provided the results were available before 5.5 hours. Infants were selected for randomization if they had a moderately or severely abnormal background aEEG voltage (moderate: upper margin of aEEG activity above 10 μV and lower margin below 5 μV; severe: upper margin below 10 μV) and/or electroencephalographically determined seizures (identified by a sudden increase in voltage accompanied by narrowing of the band of aEEG activity and followed by a brief period of suppression). Infants with aEEG seizures could be recruited despite normal or mild aEEG voltage changes, because seizures were considered to be a significant adverse prognostic factor.

Hypothermia Group

For infants assigned randomly to head cooling, a cooling cap (Olympic Medical Cool Care System; Olympic Medical, Seattle, WA) was fitted around the head for 72
hours.13–15 The system consisted of a small, thermostatically controlled, cooling unit and a pump that circulated water through the cooling cap. The initial water temperature was set to 8°C to 12°C. All infants were nursed under a radiant overhead heater, which was servo-controlled to the abdominal skin of the infant and adjusted to maintain the rectal temperature at 34.5 ± 0.5°C. Adjustments were made to the cooling cap water temperature to stay within these limits. At the time of initiation of hypothermia, the overhead heater was turned off for 20 to 30 minutes to accelerate cooling; the heater was turned back on once the rectal temperature had decreased to 35.5°C.

At the end of the 72-hour cooling period, the cooled infants were rewarmed slowly, at no more than 0.5°C per hour, until their temperature was within the normal temperature range. Cooling was discontinued earlier if the parents withdrew consent or if discontinuation was required for clinical reasons (eg, extracorporeal membrane oxygenation), in the opinion of the attending neonatologist.

Control Group
Infants assigned randomly to the control noncooled group were cared for under an overhead radiant heater, which was servo-controlled to the infants’ abdominal skin temperature to maintain the rectal temperature at 37.0 ± 0.2°C. The infants received standard clinical care for their center, under the care of the attending neonatologist.

Primary Outcome
The primary outcome was the combined incidence of death and severe neurodevelopmental disability in survivors at 18 months of age. Severe neurodevelopmental disability was defined as (1) gross motor function classification levels 3 through 5 (nonambulatory, sits with support applied to lower back, or limited or no self-mobility),16 (2) Bayley Mental Developmental Index17 of <70, or (3) bilateral cortical visual impairment. The primary outcome was analyzed according to the intent-to-treat principle. Two-sided P values of <.05 were considered statistically significant.

Exploratory Analysis
Exploratory analysis was performed by using logistic regression analyses for the primary outcome for treatment plus selected potential covariates, including grade of encephalopathy, 5-minute Apgar score, gender, birth weight, time of randomization, aEEG background, and presence of seizures determined with aEEG recording at the time of randomization, first individually and then with all covariates included in a multivariate model. We were unable to examine the specificity of changes in the arterial cord blood samples because of limited sample size for this parameter. In view of the significant effect of birth weight reported below, gestational age was also tested individually. To explore more thoroughly the effect of birth weight on outcomes, a logistic regression model with treatment/subgroup interaction terms in which birth weight was dichotomized as <25th percentile for gender at term gestation18 versus ≥25th percentile was used. Among the 218 patients, there were 72 infants (39 cooled and 33 control) in the <25th percentile group and 146 (69 cooled and 77 control) in the ≥25th percentile group. Using lower percentile cutoff points would have resulted in too few infants in the lower-weight group even for exploratory purposes. We then examined separately the hypothesis that the effect of birth weight on outcomes in the control group might be related to susceptibility to pyrexia, defined as ≥38°C in 4 hourly measurements during the observation period. Other factors, including race/ethnicity, location or number of patients treated at the sites, and maternal complications were tested separately in the main multivariate model. Outcome incidences were tested by using Fisher’s exact test.

RESULTS
Baseline and Maternal Characteristics
Outcome data at 18 months were available for 218 of 234 patients enrolled originally (93%); 8 infants from each group were lost to follow-up monitoring. Baseline patient characteristics for the groups were generally well balanced (Table 1). Because of stratification of treatment randomization according to participating site only and the generally small number of patients enrolled at each site, the Apgar score at 5 minutes after birth and aEEG background activity showed trends toward more-severe injury for infants assigned randomly to cooling. Maternal characteristics for the 218 patients with 18-month outcome data were similar to those of the total 234 patients and were generally balanced, with the exception that more control mothers experienced complications during pregnancy and labor (Table 2). These factors were tested separately in the multivariate logistic regression model (see below).

Staged Logistic Regression Analysis
In a logistic regression analysis with treatment group, the grade of encephalopathy at randomization showed the greatest predictive value. In order, greater severity of aEEG background changes, presence of aEEG seizures, lower continuous 5-minute Apgar score, and greater birth weight (in 100-g steps) were also associated with adverse outcomes (Table 3). The dichotomized Apgar score did not reach significance, although the direction of change was similar to that for the continuous Apgar score. There was no effect of gender, gestational age, or age at randomization (Table 3). The apparent effect of treatment with cooling was similar across each of these
analyses, and \( P < .05 \) after adjustment for encephalopathy grade, aEEG background, and continuous Apgar score (Table 3).

### Full Logistic Regression Model

When all of the factors were examined together in a logistic regression model (Table 4), encephalopathy grade, aEEG background, aEEG seizures, and birth weight remained independently associated with unfavorable outcomes. In contrast, Apgar score was not significantly predictive when corrected for encephalopathy grade. In this combined model, cooling was associated significantly with reduced risk of unfavorable outcomes. When this prespecified secondary analysis was repeated with only the significant variables (treatment group, encephalopathy grade, aEEG background, aEEG seizures, and birth weight), the outcome of the analysis was similar and each variable remained significantly predictive (data not shown).

### Effect of Severity of Encephalopathy

The encephalopathy grade at the time of randomization was the single most predictive covariate, both alone and in combination. There was no significant interaction between encephalopathy grade and hypothermia treatment. The incidence of unfavorable outcomes was reduced similarly after cooling for infants with either moderate (grade 2) encephalopathy (28 [45%] of 62 cooled infants and 39 [57%] of 69 control infants) or severe (grade 3) encephalopathy (28 [70%] of 40 cooled infants and 32 [91%] of 35 control infants). There were few infants with mild (grade 1) encephalopathy (2 of 5 cooled infants and 0 of 3 control infants); data were missing for 1 cooled infant and 3 control infants.

### Exploratory Analysis of Birth Weight

The analyses in Tables 3 and 4 suggested that there was a detrimental effect toward unfavorable primary outcomes with increasing weight. Infants in the control group with favorable outcomes were smaller than infants in the control group with unfavorable outcomes, weighing 3229 ± 633 g vs 3630 ± 597 g (\( P = .001 \), Mann-Whitney test), and had less-severe encephalopathy grades (2.0 ± 0.4 vs 2.5 ± 0.5; \( P = .0001 \)). To understand this effect, additional explorations were conducted in which birth weight was dichotomized. The birth weight of the infants of <25th percentile was 2785 ± 290 g, compared with 3785 ± 498 g. As shown in Table 5, there was a marked interaction between cooling and birth weight (\( P = .003 \)), as well as a significant effect of birth weight (\( P = .005 \)), controlled for severity of encephalopathy. Consistent with the previous analysis, infants in the control group with birth weights of ≥25th percentile had a significantly higher rate of unfavorable primary outcomes than did those with weights of <25th percentile (59 [77%] of 77 infants vs 14 [42%] of 33 infants; \( P < .001 \)). In contrast, there was a highly significant cooling effect in the ≥25th percentile group (unfavorable outcomes for 35 [51%] of 69 cooled infants vs 59 [77%] of 77 control infants; \( P < .002 \)) but not in the <25th percentile group (24 [62%] of 39 cooled infants vs 14 [42%] of 33 control infants; \( P = .16 \)). Acute complications in labor were reported for 124 (85%) of 146 infants of ≥25th percentile and 51 (71%) of 72
infants of <25th percentile (odds ratio [OR]: 2.32; 95% confidence interval [CI]: 1.18–4.56; \( P = .018 \), Fisher’s exact test). The number needed to treat for benefit in the ≥25th percentile group was 3.8 (95% CI: 2.5–9.6).

Pyrexia
Rectal temperature measurements confirmed that the great majority of infants in both groups achieved and maintained their target temperature ranges (Fig 1). Thirty-four control patients had rectal temperatures of ≥38°C at any time during the 76-hour monitoring period, of whom 28 had unfavorable outcomes; of the remaining 76 patients without pyrexia at any time, 45 had unfavorable outcomes (OR: 3.2; 95% CI: 1.2–8.4; \( P = .028 \), Fisher’s exact test). Only 11 patients in the hypothermia group (including 1 patient who was not cooled) experienced temperatures of ≥38°C at randomization or at rewarming, of whom 9 had unfavorable

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Maternal/Delivery Characteristics for All Patients With 18-Month Data (N = 218)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooled (N = 108)</td>
<td>Control (N = 110)</td>
</tr>
<tr>
<td>Maternal age, mean ± SD, y</td>
<td>28.7 ± 6.1</td>
</tr>
<tr>
<td>Ethnicity/race, n (%)</td>
<td></td>
</tr>
<tr>
<td>White or Asian</td>
<td>75 (69%)</td>
</tr>
<tr>
<td>Black or Maori/Pacific Islander</td>
<td>15 (14%)</td>
</tr>
<tr>
<td>Hispanic/other races</td>
<td>18 (17%)</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
</tr>
<tr>
<td>Married/stable</td>
<td>87 (81%)</td>
</tr>
<tr>
<td>Single/divorced/widowed</td>
<td>21 (19%)</td>
</tr>
<tr>
<td>Gravida, n (%)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>42 (39%)</td>
</tr>
<tr>
<td>2</td>
<td>24 (22%)</td>
</tr>
<tr>
<td>3</td>
<td>24 (22%)</td>
</tr>
<tr>
<td>4</td>
<td>8 (7%)</td>
</tr>
<tr>
<td>5</td>
<td>5 (5%)</td>
</tr>
<tr>
<td>≥6</td>
<td>4 (4%)</td>
</tr>
<tr>
<td>Parity, n (%)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>58 (54%)</td>
</tr>
<tr>
<td>1</td>
<td>25 (23%)</td>
</tr>
<tr>
<td>2</td>
<td>15 (14%)</td>
</tr>
<tr>
<td>3</td>
<td>5 (5%)</td>
</tr>
<tr>
<td>≥4</td>
<td>5 (5%)</td>
</tr>
<tr>
<td>Multiple birth, n (%)</td>
<td></td>
</tr>
<tr>
<td>6 (6%)</td>
<td>3 (3%)</td>
</tr>
<tr>
<td>Method of delivery, n (%)</td>
<td></td>
</tr>
<tr>
<td>Vaginal vertex, unassisted</td>
<td>18 (17%)</td>
</tr>
<tr>
<td>Vaginal vertex, assisted</td>
<td>10 (9%)</td>
</tr>
<tr>
<td>Vaginal breech</td>
<td>3 (3%)</td>
</tr>
<tr>
<td>Cesarean section, elective</td>
<td>2 (2%)</td>
</tr>
<tr>
<td>Cesarean section, emergency</td>
<td>75 (69%)</td>
</tr>
<tr>
<td>Delivery complications, n (%)</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>84 (78%)</td>
</tr>
<tr>
<td>Prolapsed cord</td>
<td>3 (3%)</td>
</tr>
<tr>
<td>True knot in cord</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Tear or rupture of cord</td>
<td>2 (2%)</td>
</tr>
<tr>
<td>Placental abruption</td>
<td>28 (26%)</td>
</tr>
<tr>
<td>Shoulder dystocia</td>
<td>10 (9%)</td>
</tr>
<tr>
<td>Fetomaternal bleeding</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Ruptured uterus</td>
<td>7 (6%)</td>
</tr>
<tr>
<td>Prepartum hemorrhage</td>
<td>11 (10%)</td>
</tr>
<tr>
<td>Traumatic instrument delivery</td>
<td>2 (2%)</td>
</tr>
<tr>
<td>Head entrapment</td>
<td>2 (2%)</td>
</tr>
<tr>
<td>Other delivery complications</td>
<td>47 (44%)</td>
</tr>
<tr>
<td>Maternal complications during pregnancy and labor, n (%)</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>32 (30%)</td>
</tr>
<tr>
<td>Preeclampsia or eclampsia</td>
<td>5 (5%)</td>
</tr>
<tr>
<td>Maternal seizure(s)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Antibiotic therapy</td>
<td>19 (18%)</td>
</tr>
<tr>
<td>Thyroid malfunction</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Gestational diabetes mellitus</td>
<td>7 (7%)</td>
</tr>
<tr>
<td>Ruptured membranes (&gt;18 h)</td>
<td>4 (4%)</td>
</tr>
<tr>
<td>Pyrexia of ≥37.6°C (in labor)</td>
<td>4 (4%)</td>
</tr>
<tr>
<td>Placenta previa</td>
<td>1 (1%)</td>
</tr>
</tbody>
</table>
The present secondary analysis of the CoolCap study extends recent findings that hypothermia can reduce rates of death or disability at 18 months of age after severe neonatal encephalopathy to show that outcomes of the 97 patients in the hypothermia group who did not experience pyrexia, 50 had unfavorable outcomes (OR: 4.2; 95% CI: 0.97–18; P = .11).

There was a significant but modest correlation between greater birth weight and the frequency of pyrexia of ≥38°C in 4 hourly measurements during the 76-hour monitoring period in the control group (r² = 0.05; P = .011). Both the incidence of pyrexia and birth weight (in 100-g steps) were significant when tested separately in a multivariate logistic regression model incorporating severity of encephalopathy, background aEEG changes, and presence of seizures (pyrexia: OR: 2.06; 95% CI: 1.04–4.09; P = .038; birth weight: OR: 1.1; 95% CI: 1.02–1.2; P = .012). When the incidence of pyrexia and weight were included together in the regression model, there was no significant effect of pyrexia (OR: 1.9; 95% CI: 0.9–3.7; P = .086) but birth weight remained significant (OR: 1.09; 95% CI: 1.00–1.19; P = .045).

Other Covariates
Logistic regression indicated that there was no significant interaction effect for gender (P = .16) or for Apgar score at 5 minutes (n = 213; P = .28). Maternal ethnicity/race (main effect, P = .85), delivery method (vaginal versus cesarean section; main effect, P = .9), and maternal complications during pregnancy or labor (yes versus no; main effect, P = .73) had no significant main effects or interactions with cooling treatment. Similarly, there was no effect on outcomes or interaction with treatment for the number of infants enrolled at each site (<10 vs 10–19 infants; main effect, P = .5; ≥20 vs 10–19 infants; main effect, P = .61) or non–United States versus United States sites (main effect, P = .42).

**DISCUSSION**

TABLE 3
Predictive Values for Individual Covariates Plus Treatment for Primary Outcome

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Covariate</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>OR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Encephalopathy grade, 3 vs 1 or 2</td>
<td>&lt;.0001</td>
<td>4.3 (2.21–8.38)</td>
</tr>
<tr>
<td>aEEG background, severe vs moderate</td>
<td>.0051</td>
<td>2.51 (1.23–4.78)</td>
</tr>
<tr>
<td>aEEG seizures, yes vs no</td>
<td>.011</td>
<td>2.08 (1.19–3.7)</td>
</tr>
<tr>
<td>Apgar score, continuous</td>
<td>.0052</td>
<td>0.81 (0.69–0.94)</td>
</tr>
<tr>
<td>Apgar score, &gt;4 vs 0–3</td>
<td>.279</td>
<td>0.71 (0.38–1.32)</td>
</tr>
<tr>
<td>Birth weight, 100-g steps</td>
<td>.007</td>
<td>1.07 (1.02–1.12)</td>
</tr>
<tr>
<td>Age at randomization, h</td>
<td>.82</td>
<td>1.04 (0.75–1.45)</td>
</tr>
<tr>
<td>Gender, female vs male</td>
<td>.606</td>
<td>1.1 (0.67–2.0)</td>
</tr>
<tr>
<td>Gestational age, wk</td>
<td>.166</td>
<td>1.1 (0.95–1.3)</td>
</tr>
</tbody>
</table>

Each row represents a separate binary logistic regression including the nominated covariate and the treatment group only. This analysis included all patients with primary outcome data (N = 218). The analysis used an intention-to-treat approach; 4 of the cooled infants, 2 of whom had moderate encephalopathy (grade 2) and 2 severe encephalopathy (grade 3), were not cooled.

TABLE 4
Multivariate Predictive Values for Primary Outcome

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Covariate</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>OR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Treatment, cooled vs control</td>
<td>.004</td>
<td>0.52 (0.28–0.97)</td>
</tr>
<tr>
<td>Encephalopathy grade, 3 vs 1 or 2</td>
<td>.001</td>
<td>3.37 (1.64–6.93)</td>
</tr>
<tr>
<td>aEEG background, severely abnormal vs moderately abnormal</td>
<td>.05</td>
<td>2.06 (1.01–4.17)</td>
</tr>
<tr>
<td>aEEG seizures, yes vs no</td>
<td>.04</td>
<td>1.96 (1.02–3.74)</td>
</tr>
<tr>
<td>Birth weight, 100-g steps</td>
<td>.03</td>
<td>1.06 (1.01–1.12)</td>
</tr>
<tr>
<td>Age at randomization, h</td>
<td>.75</td>
<td>1.06 (0.74–1.51)</td>
</tr>
<tr>
<td>Gender, female vs male</td>
<td>.88</td>
<td>0.95 (0.51–1.77)</td>
</tr>
<tr>
<td>Apgar score at 5 min</td>
<td>.21</td>
<td>0.90 (0.76–1.06)</td>
</tr>
</tbody>
</table>

Results are from a multivariate, binary, logistic regression model for the primary outcome (N = 218). An intention-to-treat analysis was performed; 4 of the cooled infants, 2 of whom had moderate encephalopathy (grade 2) and 2 severe encephalopathy (grade 3), were not cooled. Unknown data and minor terms such as normal/mild aEEG background activity were included in the equation as independent terms (all were nonsignificant, data not shown).

TABLE 5
Exploratory Analysis of Interaction Between Birth Weight and Treatment

<table>
<thead>
<tr>
<th>Parameter</th>
<th>P</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>.0009</td>
<td>0.28 (0.13–0.59)</td>
</tr>
<tr>
<td>Encephalopathy grade, 3 vs 1 or 2</td>
<td>&lt;.0001</td>
<td>4.13 (2.08–8.2)</td>
</tr>
<tr>
<td>Birth weight, &lt;25th vs ≥25th percentile</td>
<td>.0045</td>
<td>0.27 (0.11–0.67)</td>
</tr>
<tr>
<td>Birth weight–treatment interaction</td>
<td>.0026</td>
<td>6.67 (1.94–22.84)</td>
</tr>
</tbody>
</table>

A multivariate, binary, logistic regression analysis for the primary outcome incorporating birth weight and an interaction term for birth weight and treatment was performed (N = 218).
not only by the severity of encephalopathy but also by birth size. There was a striking interaction between treatment and weight, such that larger infants had higher rates of adverse outcomes in the control group but also a substantial apparent reduction in adverse outcomes after delayed hypothermia, compared with no significant effect of hypothermia for smaller infants. Furthermore, spontaneous pyrexia was associated with a marked increase in adverse outcomes for control infants, controlled for severity of encephalopathy. Although there was a small intriguing association between pyrexia and larger birth size, it seemed that this could not fully explain the adverse outcomes associated with greater birth weight for control infants.

This analysis was designed to explore more thoroughly patient data obtained before randomization, to identify variables that might affect outcomes in both treatment and control groups. This post hoc approach has 2 well-known limitations. First, the subset sample sizes are much smaller than the overall treatment group, which leads to an increased risk of false-negative results. Second, multiple comparisons increase the risk of false-positive findings. Nonetheless, this analysis can generate hypotheses that can be tested in ongoing clinical trials. For the purposes of this analysis, $P < .05$ was taken to indicate variables of potential future interest.

As expected, outcomes were influenced greatly by the neurologic status before randomization, with independent effects of both the severity of encephalopathy and the aEEG variables. There was a significant treatment effect, reflecting an imbalance in randomization between the groups that was reported previously. Although the 5-minute Apgar score was weakly predictive of outcomes when used by itself, the effect was not significant when corrected for encephalopathy grade, consistent with previous data, which emphasizes that the Apgar score is effectively a surrogate measure for encephalopathy. It is well established that the grade of encephalopathy determined at $\geq 24$ hours of age is closely associated with outcomes. The present findings strongly support the value of this assessment even in the first few hours of life. The incidence of adverse outcomes in the control group was similar to previous reports, although there was a slightly higher incidence of death or disability for infants with moderate encephalopathy, which probably reflects selection of more-severe cases arising from the requirement for abnormal aEEG findings. Therefore, these findings are not necessarily representative of the predictive value of encephalopathy assessments in isolation.

It is notable that our analysis suggested that hypothermia exerted neuroprotective effects in infants with both moderate and severe encephalopathy. This seems to conflict with experimental evidence suggesting that hypothermia is less effective after the most severe insults. However, a recent multicenter trial of whole-body cooling also suggested protective effects in both the moderate and severe encephalopathy groups. We speculate that some infants with severe encephalopathy had slowly evolving injuries (as reported in some experimental paradigms) and therefore retained the potential to respond to hypothermia treatment initiated within 6 hours after birth. Conversely, the presence of severe abnormalities in the aEEG recording (especially the combination of suppressed background activity and seizures) was associated with failure to respond to hypothermia, which likely reflects a subset of infants with the most-profound or most-advanced injuries. These factors are closely interrelated, because experimentally more-severe injury leads to more-rapid secondary evolution.

Unexpectedly, the present analysis indicated that birth weight was an important determinant of outcomes. In the control group, lighter infants tended to have improved outcomes, even when controlling for the severity of clinical encephalopathy in the multivariate analysis. It is important to note that this group represents the lower end of the reference range (weight below the 25th percentile for term) and not growth-restricted infants. This finding has not been reported previously, although there is some evidence that exposure to preclampsia may reduce the risk of cerebral palsy in premature infants. In experimental studies, previous exposure to hypoxia, under specific circumstances, protected against injury. It can be speculated that this so-called preconditioning might provide a mechanism through which some smaller infants might have less-severe injuries.

An alternative speculation is that, because of their greater body mass, greater thermal inertia, and lesser relative surface area, larger infants might have a greater tendency toward postasphyxial pyrexia, compared with smaller infants. It is known that even minor, delayed, induced increases in brain temperature after exposure to hypoxia/ischemia can exacerbate subsequent injury. Consistent with this possibility, there was a modest but significant correlation between birth weight and the incidence of pyrexia in 4 hourly measurements for control infants; in turn, pyrexia was associated with worse outcomes. Although more-frequent pyrexia might have contributed in part to the effect of birth weight, larger birth size was still independently associated with adverse outcomes after multivariate adjustment for severity of encephalopathy and pyrexia, which suggests that additional mechanisms are likely to be involved.

The adverse effects of pyrexia in adults with acute brain injuries have been accepted for some years, and active antipyretic therapy is widely used in adult neurointensive care. However, the CoolCap trial seems to be the first study with prospective, detailed, core temperature measurements demonstrating evidence of this association in newborn infants. The mechanisms of such spontaneous pyrexia in the noncooled infants are un-
known but likely include heat production related to intense seizures, induction of inflammatory cytokines, and possibly overshoot heating related to operation of the servo-controlled overhead heater. Because only 3 control infants had sepsis, it is likely that the adverse association was related directly to the changes in temperature. Because there has never been a therapeutic reason to make sick neonates pyrexial, our findings reinforce the importance of measuring core temperatures and preventing hyperthermia in encephalopathic infants, even if hypothermia is not used as therapy.

In the treatment group, our analysis suggested that heavier infants showed a very marked beneficial response to hypothermia, whereas lighter infants did not. The unexpected influence of birth weight on the response to treatment needs to be explored in future studies. At present, there is no support from experimental studies for this concept, and our findings may represent a statistical artifact. Because the larger infants in the control group had worse outcomes, there might simply have been more scope for improvement after hypothermia. Alternatively, the higher rate of labor complications observed for larger infants raises the possibility that the infants had a higher rate of acute peripartum insults and that the injury process was at a less-advanced stage at the time of initiation of treatment, when it was still amenable to neuroprotective intervention. In the present study, however, there was no effect of labor complications per se on the outcome of treatment.

Alternatively, the effect of infant weight might be related to cap temperatures. A recent study with 7 healthy piglets that had not received an hypoxic/ischemic insult reported that, for a given cap temperature, underlying cortical temperatures were significantly lower in smaller piglets, which raises the possibility that the brain might be overcooled in small infants. However, others showed that head cooling in newborn piglets was highly neuroprotective after hypoxia/ischemia, despite a similar level of cortical cooling. Moreover, piglets have much smaller brains than term infants, and it is unclear whether these data apply to human infants. Finally, Iwata et al examined the impact of fixed cap temperatures. However, although brain temperatures could not be measured in the CoolCap trial, the investigators’ experience is that smaller infants consistently required substantially higher (less cold) cap temperatures than larger infants. It is possible, therefore, that smaller infants might have received less-effective head cooling, with a reduced brain to rectal temperature gradient.

Male infants are known to have small increases in the incidence of encephalopathy at term, and limited data in neonatal rats raise the possibility that hypothermia might be more protective for female infants. In contrast, the present analysis suggested that there was no independent effect of gender on outcomes and no evidence of different effects of treatment for boys and girls. Furthermore, we found no effect of race, presence of maternal and labor complications, or study site on therapeutic outcomes.

Although this analysis adds to the evidence that head cooling with mild hypothermia is therapeutically valuable in neonatal encephalopathy, it is important to note that the outcomes studied to date are only from 18 months of age. Given the experimental data showing that the optimal temperatures for neuroprotection differed between the cortex and basal ganglia and the preliminary clinical finding that head cooling was associated with a greater reduction in the incidence of cortical injury, long-term follow-up monitoring is clearly needed.

The present results confirm that early encephalopathy grading is a strong independent predictor of long-term outcomes of neonatal encephalopathy, that aEEG recording provides additional predictive value, and that there is no independent effect of Apgar scores. Infants with severe and moderate encephalopathy demonstrated similar trends in responding to hypothermia treatment. Unexpectedly, in this study larger control infants had a greater incidence of death or disability even after adjustment for the effects of encephalopathy. Strikingly, however, the larger infants showed a greater therapeutic response to hypothermia. Although there was an association between greater birth weight and the incidence of pyrexia, there was an apparent independent effect of birth weight. These unexpected findings indicate important hypotheses that should be examined in current and future studies. The adverse outcomes associated with pyrexia in the present study strongly support previous recommendations that pyrexia should be rigorously prevented.

ACKNOWLEDGMENTS

The study was funded by Olympic Medical (Seattle, WA). Drug and Device Development Co (Redmond, WA) entered and held the data and performed all analyses according to the instructions of the scientific advisory committee. The sponsor supported the study financially, provided administrative support to the sites, supplied the aEEG monitors and the cooling devices, and monitored initial data recording and accuracy but had no input into the manuscript.

Children’s Hospital; 11 patients), M. Battin, D. Armstrong (University of Auckland-National Women’s Hospital, New Zealand; 11 patients), J. Khan (Children’s Memorial Hospital and Prentice Women’s Hospital of Northwestern Memorial Hospital; 3 patients), T. Raju (University of Illinois at Chicago Medical Center; 1 patient), R. Polin, R. Sahni, U. Sanocka (Children’s Hospital of New York-Presbyterian, Columbia University; 18 patients), A. Rosenberg, J. Paisley (Children’s Hospital of Denver; 23 patients), R. Goldberg, M. Cotton (Duke University Medical Center; 14 patients), A. Peliowski, E. Phillis (Royal Alexandra Hospital/University of Alberta Hospital; 20 patients), D. Azzopardi, A. D. Edwards (Hammersmith Hospital, London, United Kingdom; 1 patient), F. Northington (Johns Hopkins University; 2 patients), J. Barks, S. Donn (University of Michigan Medical Center-Mott Children’s Hospital; 12 patients), B. Couser (Children’s Hospital and Clinics of Minneapolis; 16 patients), D. Durand (Children’s Hospital and Research Center at Oakland; 8 patients), K. Sekar (Children’s Hospital of Oklahoma; 4 patients), D. Davis, M. Blayney (Children’s Hospital of Eastern Ontario/The Ottawa Hospital; 1 patient), S. Adeniyi-Jones (Al Dupont Children’s Hospital at Thomas Jefferson University Medical Center; 6 patients), T. Yanowitz (Magee Women’s Hospital/Children’s Hospital of Pittsburgh; 10 patients), R. Guillet, N. Laroia (Golisano Children’s Hospital at Strong; 10 patients), N. Finer, F. Mannino (University of California, San Diego, Medical Center Hillcrest; 8 patients), J. Partridge (University of California, San Francisco, Children’s Hospital; 2 patients), D. Davidson (Schneider Children’s Hospital; 14 patients), A. Whitelaw (Southmead Hospital, Bristol, United Kingdom; 13 patients), M. Thoresen (St Michael’s Hospital, Bristol, United Kingdom; 8 patients), J. S. Wyatt, F. O’Brien (University College Hospital, London, United Kingdom; 4 patients), B. Walsh (Vanderbilt Children’s Hospital; 13 patients), J. Perciaccante, M. O’Shea (Wake Forest University Baptist Medical Center; 1 patient); manufacturer’s representatives: J. Jones, T. Weiler, J. Mullane, D. Hammond, J. Parnell (Olympic Medical, Seattle, WA).

We thank the many technicians, nurses, physicians, and scientists at the participating sites who contributed to the development and implementation of this project and the parents who consented to enrollment of their infants in this trial; we appreciate their trust in us under conditions of great stress and anxiety. We thank the many charities and research funding agencies that supported the preliminary research necessary for this study.

REFERENCES

15. Battin MR, Penrice J, Gunn TR, Gunn AJ. Treatment of term infants with head cooling and mild systemic hypothermia (35.0°C and 34.5°C) after perinatal asphyxia. Pediatrics. 2003;111:244–251
17. Bayley N. Bayley Scales of Infant Development. 2nd ed. San Antonio, TX: Psychological Corp; 1993

TIME-ZONED: WORKING AROUND THE ROUND-THE-CLOCK WORKDAY

“[There is a] growing army of employees whose work spans so many time zones that 18-hour workdays are routine. The long hours pose not only the risk of burnout and lost creativity, but ripple out to touch all facets of employees’ lives. A friend who is a globetrotting manager is remarkable for her unswerving emotional presence with people she cares about—except when her direct reports in China call during dinner. Then, with apologies, she has to turn her attention to business, interrupting our time together. Conventional approaches to flexibility, such as flextime, don’t help these time-zone warriors.”


Noted by JFL, MD
Mechanisms, Clinical Presentations, Injuries, and Outcomes From Inflicted Versus Noninflicted Head Trauma During Infancy: Results of a Prospective, Multicentered, Comparative Study

Kent P. Hymel, MD*a,b, Kathi L. Makoroff, MD*c, Antoinette L. Laskey, MD, MPH*d, Mark R. Conaway, PhD*e, James A. Blackman, MD, MPH*a

*aDepartment of Pediatrics and *Division of Biostatistics and Epidemiology, University of Virginia, Charlottesville, Virginia; bInova Fairfax Hospital for Children, Falls Church, Virginia; cDepartment of Pediatrics, University of Cincinnati College of Medicine, Mayerson Center for Safe and Healthy Children, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio; dDepartment of Pediatrics, Indiana University School of Medicine, Indianapolis, Indiana

DOI:10.1542/peds.2006-3111

ABSTRACT

OBJECTIVE. Our goal was to conduct a prospective, multicentered, comparative study that would objectively verify and explain observed differences in short-term neurodevelopmental outcomes after inflicted versus noninflicted head trauma.

METHODS. Children <36 months of age who were hospitalized with acute head trauma confirmed by computed tomography imaging were recruited at multiple sites. Extensive clinical data were captured prospectively, subjects were examined, cranial imaging studies were blindly reviewed, and caregivers underwent scripted interviews. Follow-up neurodevelopmental evaluations were completed 6 months after injury. Head-trauma etiology and mechanisms were categorized by using objective a priori criteria. Thereafter, subject groups with inflicted versus noninflicted etiologies were compared.

RESULTS. Fifty-four subjects who met the eligibility criteria were enrolled at 9 sites. Of 52 surviving subjects, 27 underwent follow-up assessment 6 months after injury. Etiology was categorized as noninflicted in 30 subjects, inflicted in 11, and undetermined in 13. Compared with subjects with noninflicted head trauma, subjects with inflicted head trauma (1) more frequently experienced noncontact injury mechanisms, (2) sustained greater injury depth, (3) more frequently manifested acute cardiorespiratory compromise, (4) had lower initial Glasgow Coma Scale scores, (5) experienced more frequent and prolonged impairments of consciousness, (6) more frequently demonstrated bilateral, hypoxic-ischemic brain injury, (7) had lower mental developmental index scores 6 months postinjury, and (8) had lower gross motor quotient scores 6 months postinjury.

CONCLUSIONS. Compared with infants with noninflicted head trauma, young victims of inflicted head trauma experience more frequent noncontact injury mechanisms that result in deeper brain injuries, cardiorespiratory compromise, diffuse cerebral hypoxia-ischemia, and worse outcomes.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-3111

doi:10.1542/peds.2006-3111

Key Words
head trauma, infants, child abuse, accidents, outcomes

Abbreviations
GCS—Glasgow Coma Scale
PediBIRN—Pediatric Brain Injury Research Network
CT—computed tomography
MDI—mental developmental index
GMQ— gross motor quotient
MVC—motor vehicle crash
SDH—subdural hematoma

Accepted for publication Jan 16, 2007

Address correspondence Kent P. Hymel, MD, Inova Fairfax Hospital for Children, Department of Pediatrics, 3300 Gallows Rd, Falls Church, VA 22042-3300. E-mail: kent.hymel@inova.com

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
During infancy, inflicted or abusive head trauma is a leading cause of traumatic death and disability, accounting for nearly one quarter of all pediatric hospital admissions for head injury and two thirds of infant homicides. Many survivors of inflicted head trauma during infancy later manifest developmental delays, sensory deficits, feeding difficulties, recurrent seizures, intellectual deficits, motor impairments, attention deficits, and/or educational and behavioral dysfunctions.

Published studies that compare outcomes after inflicted versus noninflicted pediatric head trauma are limited in number, largely retrospective, and use a wide variety of definitional criteria for inflicted versus noninflicted etiologies. Some of these definitional criteria are fundamentally flawed, incorporating inappropriate circular logic or inherent biases. Despite their limitations, these studies provide compelling evidence that outcomes after inflicted head trauma are generally worse than outcomes after noninflicted pediatric head trauma.

Abusive acts have been linked to subdural, subarachnoid, and retinal hemorrhaging; localized axonal injury in the region of the craniocervical junction and cervical cord; acute respiratory compromise or arrest; loss of consciousness; hypotension; and secondary, diffuse, hypoxic-ischemic brain injury with swelling. We suspect that these primary and secondary inflicted traumatic injuries and their clinical manifestations are best explained by noncontact injury mechanisms (ie, rotational cranial acceleration and/or deceleration). Therefore, we hypothesized that compared with infants with noninflicted head trauma, infants with inflicted head trauma experience more frequent noncontact injury mechanisms, greater injury depth on neuroimaging, more frequent acute cardiorespiratory compromise, lower initial Glasgow Coma Scale (GCS) scores, more frequent and prolonged impairments of consciousness, more frequent diffuse, hypoxic-ischemic brain injuries, and worse short-term outcomes.

Our objective was to apply widely acceptable, a priori definitional criteria for etiology and mechanism(s) of injury in a prospective, multicentered, comparative study that would verify and explain observed differences in short-term, neurodevelopmental outcomes after inflicted versus noninflicted head trauma during infancy.

METHODS
The Pediatric Brain Injury Research Network (PediBIRN) is a consortium of clinical investigators who have made a voluntary commitment to conduct collaborative, multicentered, clinical research regarding traumatic brain injuries in young children. The institutional review boards at all 9 participating PediBIRN institutions approved this research study before local subject recruitment. Funded research activities began in March 2003 at the University of Virginia and at Inova Fairfax Hospital for Children. The remaining unfunded PediBIRN sites began subject recruitment at various times thereafter, enrolling a convenience sample of their eligible subjects.

Participants
At every site, hospital admission logs and medical charts were screened to identify eligible study subjects. Inclusion criteria included age <36 months and computed tomography (CT) evidence of any acute, nonpenetrating head or brain injury leading to inpatient evaluation and/or treatment. Children with preexisting brain disease, infection, hypoxia-ischemia, or trauma; birth injury; developmental delays; sensory deficits; bleeding disorder or severe malnutrition were excluded from study participation. Screening for exclusion criteria was accomplished through scripted parental interviews, interpretation of growth parameters, and subsequent blinded reviews of the subjects’ initial cranial CT scans.

Procedures
Subjects who met inclusion and exclusion criteria and whose parent(s) consented to study participation were examined serially during the course of their acute hospitalizations. Local researchers reviewed their emergency medical technician, emergency department, and inpatient medical charts to capture extensive demographic, historical, clinical, laboratory, and neurosurgical data. The subjects’ primary caregivers (the persons responsible for these children when they each became clearly and persistently ill with clinical signs later linked to their acute, traumatic, cranial injuries) were interviewed extensively by using a scripted interview designed to capture essential historical and clinical data from the scene of injury in a consistent and objective manner. More specifically, every primary caregiver who consented to be interviewed was questioned systematically regarding the specific timing, etiology (inflicted versus noninflicted), and circumstances of injury; the infant’s clinical and mental status before and after his/her acute deterioration; and medical interventions at the scene of injury. Qualified neuroradiologists blinded to all other clinical and historical information reviewed the subjects’ complete CT and MRI cranial imaging studies to document injuries and to identify the greatest depth of visible injury. Whenever possible, child protection and/or police investigators were interviewed to capture additional historical and investigative information.

Follow-up Neurodevelopmental Assessments
The surviving subjects returned 6 months after injury for an extensive, outpatient, neurodevelopmental assessment that included the Bayley Scales of Infant Development (second edition) and the Peabody Developmental Motor Scales (second edition). The fully qualified physical or occupational therapists, rehabilitation specialists, or developmental pediatricians who conducted these assessments at each participating site were blinded to all
historical, clinical, and radiologic information. The mental developmental index (MDI) and gross motor quotient (GMQ) scores calculated during these follow-up assessments were used as outcome measures of overall cognitive and motor development, respectively. Both are standardized scores like an IQ, with mean values of 100.

**Statistical Analyses**

All data were captured on data forms that were piloted extensively at the 2 funded PediBIRN institutions. Subsequently, the complete data were transferred into a relational database (Access 2000; Microsoft, Inc, Redmond, WA) designed to facilitate subsequent analyses. To enhance objectivity, database queries were created to sort subjects according to specific, a priori, definitional criteria for etiology (Table 1) and mechanism(s) of injury (Table 2). Other queries were designed to facilitate determinations regarding injury depth and the duration of impaired consciousness. Finally, subject groups were compared and outcome data were analyzed in an attempt to identify predictors of MDI and GMQ scores 6 months after injury.

χ² tests were used to compare outcome variables, such as mechanisms of injury, the duration of impaired consciousness, and depth of injury among subject groups defined by etiology, categorized as inflicted, noninflicted, or undetermined. The nonparametric Kruskal-Wallis test was used to compare continuous outcome variables such as the 6-month MDI and GMQ scores among groups defined by etiology, mechanism of injury, or depth of injury. Spearman rank correlations were used to assess the associations between continuous variables. Analyses were conducted in SAS 9.1 (SAS Institute, Inc, Cary, NC) and GAUSS 6.1 (Aptech Systems, Inc, Black Diamond, WA).

**RESULTS**

Twenty-seven eligible subjects were recruited at the University of Virginia or at Inova Fairfax Hospital for Children over a 2½-year period between 2003 and 2006. The remaining PediBIRN sites enrolled a convenience sample of 27 additional subjects during the same time interval. Overall, 54 study subjects meeting eligibility criteria were enrolled.

Etiology was categorized as inflicted in 11, undetermined in 13, and noninflicted in 30 subjects. Among subjects meeting a priori criteria for noninflicted head trauma were 4 children head-injured in a motor vehicle collision and 26 whose injuries were attributed to a fall. These falls included 4 involving stairs, 3 from heights >10 feet, and 2 from heights between 6 and 10 feet. The remaining subjects reportedly fell from heights <6 feet.

Demographic, mechanistic, clinical, neuroimaging, and outcome data regarding our comparative subject groups are summarized in Table 3. Age at the time of injury, gender, and racial and ethnic distributions were similar in subject groups meeting criteria for inflicted and noninflicted head trauma. Compared with our 30 subjects with noninflicted head injuries, the 11 subjects with inflicted head injuries (1) more frequently experienced noncontact injury mechanisms (P < .001), (2) revealed greater depth of injury on neuroimaging (P < .001), (3) more frequently manifested signs of acute cardiorespiratory compromise (P < .001), (4) had lower

### Table 1: A Priori Criteria for Categorizing the Etiology of Head Injuries

<table>
<thead>
<tr>
<th>Category</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noninflicted</td>
<td>Cases in which the child's primary caregiver described an accidental head injury event that was developmentally consistent, historically consistent with repetition over time, could be linked to the child's acute clinical presentation for traumatic cranial injuries, and occurred in the absence of any noncranial injuries considered moderately or highly specific for abuse*</td>
</tr>
<tr>
<td>Inflicted</td>
<td>Cases in which the child's primary caregiver admitted abusive acts that could be linked to the child's acute clinical presentation for traumatic cranial injuries</td>
</tr>
<tr>
<td></td>
<td>Cases in which an independent witness verified abusive acts that could be linked to the child's acute clinical presentation for traumatic cranial injuries</td>
</tr>
<tr>
<td></td>
<td>Cases in which a child not yet cruising or walking became clearly and persistently ill with signs of acute cardiorespiratory compromiseb linked to his/her traumatic cranial injuries while in the care of a primary caregiver who denied any knowledge of a head injury event</td>
</tr>
<tr>
<td>Undetermined</td>
<td>Cases in which the child's primary caregiver provided an explanation for the child's head injury event that was clearly developmentally inconsistent with the parent(s)' description of their child's developmental capabilities</td>
</tr>
<tr>
<td></td>
<td>Cases in which an independent witness verified abusive acts that could be linked to the child's acute clinical presentation for traumatic cranial injuries</td>
</tr>
<tr>
<td></td>
<td>Cases in which a child not yet cruising or walking became clearly and persistently ill with signs of acute cardiorespiratory compromiseb linked to his/her traumatic cranial injuries while in the care of a primary caregiver who denied any knowledge of a head injury event</td>
</tr>
</tbody>
</table>

*Including classic metaphyseal lesion(s), fractures of the ribs, scapula, sternum, spinous process(es), or digit(s); vertebral body fracture(s) or dislocation; epiphyseal separation(s); noncranial bruising, abrasion(s) or laceration(s) in location(s) other than the knees, shins or elbows; patterned bruise(s) or dry contact burn(s); scaling burns with uniform depth, clear lines of demarcation, and a paucity of splash marks; intra-abdominal injuries; retinal hemorrhages described by an ophthalmologist as dense, extensive, covering a large surface area of the retina, or extending to the periphery of the retina; and retinoschisis diagnosed by an ophthalmologist.

b Including breathing difficulty, respiratory distress, infrequent respirations, apnea, or cyanosis; clinical manifestations of shock, delayed capillary refill, or cardiac arrest; any requirement for mouth-to-mouth breathing, bag-mask ventilation, intubation, chest compressions, rapid volume expansion; or epinephrine therapy; occurring at the scene of injury, during transport, in the emergency department, or at the time of hospital admission; documented by medical personnel or reported by the child's primary caregiver.
Historical multivariable analyses. too few subjects with follow-up data to conduct additional multivariable analyses.

TABLE 2 | A Priori Criteria for Categorizing the Mechanism(s) of Head Injury

<table>
<thead>
<tr>
<th>Category</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact</td>
<td>Cases with injuries limited to craniofacial bruising, abrasion(s), laceration(s) or swelling; subgaleal hematoma(s); cephalohematoma(s); skull fracture(s) or epidural hematoma</td>
</tr>
<tr>
<td>Noncontact</td>
<td>Cases with injuries limited to acute concussion, diffuse axonal injury, or an abnormal subdural collection that extends from the interhemispheric region</td>
</tr>
<tr>
<td>Combined</td>
<td>Cases with both contact and noncontact injuries</td>
</tr>
<tr>
<td>Undetermined</td>
<td>Cases with injuries limited to subarachnoid hemorrhage, brain contusion(s), brain laceration(s), or an abnormal subdural collection that does not extend from the interhemispheric region</td>
</tr>
</tbody>
</table>

*a* Contact injuries can be viewed as the deformingal injuries that result from cranial impact if the head is prevented from moving. Primary brain injuries resulting from isolated contact mechanisms result solely from cerebral tissue distortions induced by skull deformation.  
*b* Noncontact injuries result solely from cranial acceleration or deceleration, irrespective of whether or not there is a direct impact to the cranium.  
*c* These undetermined injuries can result from either contact or noncontact injury mechanisms.

Initial GCS scores \(P < .001\), (5) experienced more frequent and prolonged impairments of consciousness \(P < .001\), (6) more frequently demonstrated bilateral, hypoxic-ischemic, brain injury, or swelling \(P < .001\), (7) had lower MDI scores 6 months postinjury \(P = .011\), depth of injury \(P = .003\), signs of acute cardiorespiratory compromise \(P = .227\), and initial GCS scores \(P = .001\). Their GMQ scores 6 months postinjury were related significantly to etiology \(P = .002\), mechanisms of injury \(P = .05\), depth of injury \(P = .01\), signs of acute cardiorespiratory compromise \(P = .005\), the duration of impaired consciousness \(P = .013\), and initial GCS scores \(P < .001\).

Among the surviving subjects meeting definitional criteria for noninflicted head trauma, none who were evaluated 6 months postinjury had an MDI or GMQ score <60. In contrast, among the subjects meeting definitional criteria for inflicted head trauma, 3 (75%) of 4 who underwent a Bayley assessment had MDI scores <60, and 4 (67%) of 6 who underwent a Peabody assessment had GMQ scores <60. Our study included too few subjects with follow-up data to conduct additional multivariable analyses.

**DISCUSSION**

Poor outcomes after pediatric head trauma have been linked to low initial GCS scores, increased depth and duration of impaired consciousness, diffuse cerebral edema, cerebral hypoperfusion, brain infarction, and increased depth of parenchymal injury.\(^5,24,25,27-31\) Our results confirm specifically all of these conclusions from previous studies (Tables 3 and 4).

Neuropathological studies of abused infants have identified localized injuries in the region of the cranio-cervical junction that can trigger acute concussion, apnea, respiratory insufficiency, bradycardia, shock, or cardiac arrest.\(^20-22\) We suspect that these deep injuries result from noncontact (that is, rotational cranial acceleration or deceleration) injury mechanisms resulting from abusive acts. If not promptly resolved or treated, the resulting cardiopulmonary compromise will initiate or accelerate the development of secondary, diffuse, hypoxic-ischemic brain injuries that have been linked to poor outcomes. Through these pathophysiological processes, etiology significantly influences outcome.

On the basis of these considerations, we hypothesized that, compared with infants with noninflicted head trauma, infants with inflicted head trauma experience more frequent noncontact injury mechanisms, greater injury depth, more frequent signs of acute cardiorespiratory compromise, lower initial GCS scores, more frequent and prolonged impairments of consciousness, more frequent diffuse hypoxic-ischemic brain injuries, and worse short-term outcomes. Our results confirm these hypotheses (Table 3) and provide preliminary evidence that young victims of inflicted head trauma have poor outcomes because they frequently experience noncontact injury mechanisms.

Compared with our subjects with noninflicted head injuries, our surviving subjects with inflicted head trauma revealed MDI and GMQ scores 6 months postinjury that were significantly and markedly lower. In light of so many other significant differences summarized earlier, this result is not surprising. Although there are limitations on the use of MDI scores and other Bayley results for prediction of future cognitive function, such large discrepancies suggest that clinicians should be cautious regarding prognosis after inflicted head trauma, and perhaps more optimistic after noninflicted head trauma. The same could be said for motor prognosis, even more confidently, as evaluation of motor skills in this age group is more reliable.

Controversy persists in the medical literature regarding the dangers of shaking an infant. Although published case reports and series have attributed pediatric intracranial injuries to isolated shaking,\(^32,33\) some authors contend that shaking is not injurious and/or that perpetrator admissions of shaking are unreliable.\(^34,35\) In this multicentered study, we used a scripted interview to capture historical data from caregivers in a consistent manner. None of the caregivers interviewed during this study disclosed abusive acts. Nevertheless, it is interesting to note that 8 of 11 subjects meeting a priori criteria for inflicted pediatric head trauma (Table 1) manifested cranial injuries indicative of an isolated noncontact mechanism of injury.
### TABLE 3  Demographic, Mechanistic, Clinical, Neuroimaging, and Outcome Data of Comparative Subject Groups

<table>
<thead>
<tr>
<th></th>
<th>Inflicted (N = 11)</th>
<th>Noninflicted (N = 30)</th>
<th>Undetermined (N = 13)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at the time of injury, mo</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>5</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>10.5</td>
<td>9</td>
<td>11.5</td>
</tr>
<tr>
<td>SD (range)</td>
<td>10 (0.5–27)</td>
<td>10 (0.5–28)</td>
<td>11 (2–31)</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>5 (45)</td>
<td>14 (47)</td>
<td>7 (54)</td>
</tr>
<tr>
<td>Male</td>
<td>6 (55)</td>
<td>16 (53)</td>
<td>6 (46)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>8 (73)</td>
<td>27 (90)</td>
<td>7 (54)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>0 (0)</td>
<td>2 (7)</td>
<td>2 (15)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (27)</td>
<td>1 (3)</td>
<td>4 (31)</td>
</tr>
<tr>
<td><strong>Ethnicity, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>3 (27)</td>
<td>2 (7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Non-Hispanic, non-Latino, or unknown ethnicity</td>
<td>8 (73)</td>
<td>28 (93)</td>
<td>13 (100)</td>
</tr>
<tr>
<td>**Mechanism(s) of injury, n (%)**a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isolated contact mechanism</td>
<td>1 (3)</td>
<td>19 (70)</td>
<td>8 (62)</td>
</tr>
<tr>
<td>Isolated noncontact mechanism</td>
<td>8 (73)</td>
<td>1 (4)</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Combined contact and noncontact mechanism(s)</td>
<td>2 (18)</td>
<td>7 (26)</td>
<td>3 (23)</td>
</tr>
<tr>
<td>Undetermined mechanism(s)</td>
<td>0 (0)</td>
<td>3 (10)</td>
<td>1 (8)</td>
</tr>
<tr>
<td>**Greatest depth of injuries visible on neuroimaging, n (%)**a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scalp</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Skull</td>
<td>0 (0)</td>
<td>13 (43)</td>
<td>3 (23)</td>
</tr>
<tr>
<td>Epidural space</td>
<td>0 (0)</td>
<td>1 (3)</td>
<td>2 (15)</td>
</tr>
<tr>
<td>Extra-axial spaces</td>
<td>1 (9)</td>
<td>8 (27)</td>
<td>4 (31)</td>
</tr>
<tr>
<td>Cortical brain</td>
<td>2 (18)</td>
<td>6 (20)</td>
<td>2 (15)</td>
</tr>
<tr>
<td>Subcortical brain</td>
<td>8 (73)</td>
<td>2 (7)</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Brain stem or cerebellum</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>**Signs of acute cardiorespiratory compromise, n (%)**b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>4 (36)</td>
<td>27 (90)</td>
<td>12 (92)</td>
</tr>
<tr>
<td>Present</td>
<td>7 (64)</td>
<td>3 (10)</td>
<td>1 (8)</td>
</tr>
<tr>
<td><strong>Initial GCS</strong>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>11</td>
<td>29</td>
<td>13</td>
</tr>
<tr>
<td>Mean</td>
<td>8.0</td>
<td>13.2</td>
<td>13.7</td>
</tr>
<tr>
<td>SD</td>
<td>4.4</td>
<td>2.9</td>
<td>13.7</td>
</tr>
<tr>
<td>**Duration of impaired consciousness, n (%)**a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None reported or documented</td>
<td>0 (0)</td>
<td>24 (80)</td>
<td>11 (85)</td>
</tr>
<tr>
<td>Only at the scene of injury</td>
<td>5 (45)</td>
<td>3 (10)</td>
<td>1 (8)</td>
</tr>
<tr>
<td>At hospital admission but persistent for &lt;6 h</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>6–24 h, responsive by 24 h</td>
<td>1 (9)</td>
<td>1 (3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>&gt;24 h, never decerebrate, decorticate, or flaccid</td>
<td>1 (9)</td>
<td>1 (3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>&gt;24 h, decerebrate, decorticate, or flaccid</td>
<td>4 (36)</td>
<td>1 (3)</td>
<td>1 (8)</td>
</tr>
<tr>
<td><strong>Traumatic injuries visible on neuroimaging, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soft tissue injuries</td>
<td>3 (27)</td>
<td>23 (77)</td>
<td>9 (69)</td>
</tr>
<tr>
<td>Skull fracture(s)</td>
<td>2 (18)</td>
<td>23 (77)</td>
<td>8 (62)</td>
</tr>
<tr>
<td>Epidural hematoma</td>
<td>1 (9)</td>
<td>2 (7)</td>
<td>2 (15)</td>
</tr>
<tr>
<td>Subdural hemorrhage</td>
<td>8 (73)</td>
<td>9 (30)</td>
<td>3 (23)</td>
</tr>
<tr>
<td>Subarachnoid hemorrhage</td>
<td>3 (27)</td>
<td>9 (30)</td>
<td>5 (38)</td>
</tr>
<tr>
<td>Brain contusion(s) or laceration(s)</td>
<td>7 (64)</td>
<td>8 (27)</td>
<td>2 (15)</td>
</tr>
<tr>
<td>Diffuse axonal injury</td>
<td>3 (27)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Unilateral hypoxia, ischemia, or swelling</td>
<td>2 (18)</td>
<td>2 (7)</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Bilateral hypoxia, ischemia, or swellinga</td>
<td>7 (64)</td>
<td>1 (3)</td>
<td>1 (8)</td>
</tr>
<tr>
<td><strong>Death from head injuries, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>10 (91)</td>
<td>29 (97)</td>
<td>13 (100)</td>
</tr>
<tr>
<td>Yes</td>
<td>1 (9)</td>
<td>1 (3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>MDI standardized score 6 mo postinjury</strong>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>4</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>60.0</td>
<td>94.4</td>
<td>107.3</td>
</tr>
<tr>
<td>SD</td>
<td>20.0</td>
<td>11.8</td>
<td>11.7</td>
</tr>
<tr>
<td><strong>GMQ standardized score 6 mo postinjury</strong>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>6</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>59.8</td>
<td>101.8</td>
<td>102.2</td>
</tr>
<tr>
<td>SD</td>
<td>21.5</td>
<td>10.2</td>
<td>6.9</td>
</tr>
</tbody>
</table>

a Observed differences between inflicted (n = 11) and noninflicted (n = 30) subject groups were highly significant with P < .001 by χ² testing.

b Including breathing difficulty, respiratory distress, infrequent respirations, apneal or cyanosis; clinical manifestations of shock, delayed capillary refill, or cardiac arrest; any requirement for mouth-to-mouth breathing, bag-mask ventilation, intubation, chest compressions, rapid volume expansion, or epinephrine therapy; occurring at the scene of injury, during transport, in the emergency department, or at the time of hospital admission; documented by medical personnel or reported by the child’s primary caregiver.

c Observed differences between inflicted (n = 11) and noninflicted (n = 30) subject groups were significant with P < .05 by Kruskal-Wallis testing.
The 30 subjects who met definitional criteria for non-inflicted head injury included 4 very young children injured during a motor vehicle crash (MVC). Two of these subjects were not properly restrained at the time, including a 1-month-old who was in his mother’s arms and a 7-month-old restrained in a car seat that was facing forward rather than backward. Both presented with acute, life-threatening clinical signs and with alterations of consciousness lasting \(24\) hours. Both experienced noncontact injury mechanisms—1 in isolation and the other with associated contact injuries. Both demonstrated deep brain injuries involving their subcortical brains. The 1-month-old infant died, the only fatality among our 30 subjects with noninflicted head injury. In contrast, the other 2 young children who sustained their cranial injuries during a MVC were still properly restrained facing backward in their car seats after their MVC. Both sustained a skull fracture without any visible intracranial or brain injuries. Neither experienced an impairment of consciousness.

A large body of literature supports a conclusion that short-distance, pediatric falls are generally benign.\(^3_6\)–\(^4_1\). Our results also support this conclusion. Six months after injury, our 26 subjects with noninflicted head injury attributed to a fall demonstrated normal MDI and GMQ scores (mean: 93.7 [SD: 11.8] and 102 [SD: 10.8], respectively). On the other hand, our data revealed a few interesting exceptions.

A 1½-month-old black child fell 6 to 10 feet onto concrete. His caregiver heard him fall, responded rapidly, and reported that he cried immediately but was readily consolable. Thirty minutes later, he developed an alteration of consciousness lasting \(6\) hours and was associated with brief periods of flaccidity, unresponsiveness, seizure activity, and apnea without cyanosis. His cranial CT scan revealed only skull fracture without any visible, underlying intracranial or brain injury. He improved dramatically and was discharged on his second hospital day, but did not return 6 months postinjury for a follow-up assessment. This case seems to demonstrate

### TABLE 4  Predictors of Outcome 6 Months After Pediatric Head Trauma

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Etiology of injury(^a)</th>
<th>Mechanism(s) of injury(^b)</th>
<th>Greatest depth of injuries visible on neuroimaging</th>
<th>Signs of acute cardiorespiratory compromise(^c)</th>
<th>Duration of impaired consciousness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Isolated contact mechanism</td>
<td>Scalp</td>
<td>Absent</td>
<td>None reported or documented</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Skull</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Epideral space</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Extra-axial spaces</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cortical brain</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Subcortical brain</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Brain stem or cerebellum</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predictor</th>
<th>MDI Scores 6 mo Postinjury</th>
<th>GMQ Scores 6 mo Postinjury</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Etiology of injury(^a)</td>
<td>94.4</td>
<td>11.8</td>
</tr>
<tr>
<td>Noninflicted</td>
<td>60.0</td>
<td>20.0</td>
</tr>
<tr>
<td>Inflicted</td>
<td>94.8</td>
<td>11.8</td>
</tr>
<tr>
<td>Mechanism(s) of injury(^b)</td>
<td>94.8</td>
<td>11.8</td>
</tr>
<tr>
<td>Isolated contact mechanism</td>
<td>94.8</td>
<td>11.8</td>
</tr>
<tr>
<td>Isolated noncontact mechanism</td>
<td>73.0</td>
<td>39.8</td>
</tr>
<tr>
<td>Combined contact and noncontact mechanism(s)</td>
<td>90.5</td>
<td>30.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Greatest depth of injuries visible on neuroimaging</th>
<th>Scalp</th>
<th>Skull</th>
<th>Epideral space</th>
<th>Extra-axial spaces</th>
<th>Cortical brain</th>
<th>Subcortical brain</th>
<th>Brain stem or cerebellum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>—</td>
<td>94.1</td>
<td>92.0</td>
<td>110.0</td>
<td>89.5</td>
<td>50.0</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.3</td>
<td>9.9</td>
<td>9.0</td>
<td>10.8</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initial GCS score</th>
<th>Correlation</th>
<th>95% CI</th>
<th>N</th>
<th>P</th>
<th>Correlation</th>
<th>95% CI</th>
<th>N</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.64</td>
<td>(0.32–0.83)</td>
<td>26</td>
<td>.001</td>
<td>0.74</td>
<td>(0.46–0.89)</td>
<td>22</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Ci indicates confidence interval.

\(a\) See Table 1 for specific, prospective, definitional criteria.

\(b\) See Table 2 for specific, prospective, definitional criteria.

\(c\) Including breathing difficulty, respiratory distress, infrequent respirations, apnea, or cyanosis; clinical manifestations of shock, delayed capillary refill, or cardiac arrest; any requirement for mouth-to-mouth breathing, bag-mask ventilation, intubation, chest compressions, rapid volume expansion, or epinephrine therapy; occurring at the scene of injury, during transport, in the emergency department, or at the time of hospital admission; documented by medical personnel or reported by the child’s primary caregiver.
that delayed and significant clinical deterioration can occur after closed pediatric head trauma, even in the absence of visible intracranial injuries.

Four of our 26 subjects with noninflicted head injuries attributed to a fall manifested an impairment of consciousness. Seven of 26 revealed subdural hematoma (SDH) on neuroimaging studies. These 7 subjects included 3 with SDH that extended or originated from the interhemispheric region, indicative of a noncontact mechanism of injury. All 3 of these young children fell from a height of only 3 to 6 feet; manifested associated contact injuries; revealed only superficial, focal, cortical brain injuries; and experienced no alterations of consciousness. These cases seem to demonstrate that noncontact (ie, interhemispheric) SDH of minimal clinical significance can result from a short-distance fall and that the injury thresholds or biomechanical requirements for noncontact SDH and acute concussion, also a noncontact injury, are different.

Our study’s primary limitations include its small sample size, a potential for sampling bias, and the limited number of subjects with follow-up neurodevelopmental assessments. Despite these limitations, our data revealed highly significant differences between comparison groups. With additional MDI and GMQ scores, we could have conducted the appropriate multivariable analyses to identify which variables are independent and significant predictors of neurodevelopmental outcome 6 months after head injury in this young age group.

In previously published comparative studies, a wide variety of research criteria were used to differentiate between inflicted and noninflicted pediatric head trauma. Some of these definitional criteria are fundamentally flawed. By incorporating elements that refer to specific cranial injuries, some authors introduced inappropriate circular logic (for example, many young victims of abuse manifest subdural bleeding. Therefore, this child with subdural bleeding must have been abused.) By incorporating definitional elements that compared a child’s specific cranial injuries to the caregiver’s specific explanation for those injuries, some authors introduced inappropriate biases (eg, Short falls do not cause subdural bleeding).

To overcome these deficiencies, in this prospective comparative study, we categorized the etiology of each subject’s head trauma as inflicted, noninflicted, or undetermined by using the a priori criteria summarized in Table 1. These criteria incorporate the results of a recent physician survey designed to identify widely acceptable definitional criteria for inflicted and noninflicted pediatric head trauma. Furthermore, these criteria judiciously avoid circular logic and other inherent biases. For example, in the absence of any other specific criteria for inflicted or noninflicted etiology, if the child’s primary caregiver at the time of injury could not be interviewed, then cases with consistent documentation of an accidental head injury event were nevertheless classified as undetermined with respect to etiology. Similarly, in the absence of any other specific criteria for inflicted or noninflicted etiology, cases with a single, isolated, noncranial injury considered moderately or highly specific for abuse (see Table 1, footnote 1) were also classified as undetermined with respect to etiology.

Our study’s primary strengths include its prospective, multicentered design and the a priori application of widely acceptable, definitional criteria for inflicted versus noninflicted etiology that are free of circular logic and other inherent biases. Applied to our data set, our definitional criteria for etiology categorized more cases as undetermined (n = 13) than inflicted (n = 11). This result serves to reinforce our impression that our a priori criteria for etiology are appropriately objective and conservative.

CONCLUSIONS

Compared with infants with noninflicted head trauma, young victims of inflicted head trauma experience more frequent noncontact injury mechanisms that can trigger deeper brain injuries, acute cardiorespiratory compromise, secondary diffuse cerebral hypoxia-ischemia, and worse outcomes. During infancy, etiology is a significant predictor of short-term neurodevelopmental outcome after head injury. Young victims of inflicted head trauma require thorough neurodevelopmental assessment and monitoring.

ACKNOWLEDGMENTS

This work was supported, in part, with resources provided by the National Institutes of Health (grants R21HD043351-01 and R24HD39361-05), the American Academy for Cerebral Palsy and Developmental Medicine, the University of Virginia Children’s Hospital Research Fund, and PediBIRN.

In addition to the authors, the following PediBIRN participants and sites also made substantial contributions to research project design and acquisition of data: Michael A. Stoiko, MD, and Amy S. Coombs, RN (Michigan State University); Deborah E. Lowen, MD (University of Oklahoma); Nancy S. Harper, MD, and Carole Jenny, MD, MBA (Brown University); Bruce E. Herman, MD (University of Utah); and Suzanne M. Edmunds, MD (Wake Forest University). We thank Patrick Hartsfield and Teresa Olsen for vital assistance with data entry and management.

REFERENCES

3. Overpeck MD, Brenner RA, Trumble AC, Trifiletti LB, Beren-


Developmental Outcome After Epilepsy Surgery in Infancy

Tobias Loddenkemper, MD, Katherine D. Holland, MD, PhD, Lisa D. Stanford, PhD, Prakash Kotagal, MD, William Bingaman, MD, Elaine Wyllie, MD

Departments of Pediatric Neurology and Pediatric Neurosurgery, Cleveland Clinic, Cleveland, Ohio; Division of Neurology, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio; Department of Psychiatry, University of Illinois College of Medicine, Chicago, Illinois

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

ABSTRACT

OBJECTIVES. Our goals were to determine the effect of epilepsy surgery in infants (<3 years of age) on development and describe factors associated with postoperative developmental outcome.

METHODS. We identified 50 infants among 251 consecutive pediatric patients (<18 years old) undergoing epilepsy surgery. Charts were reviewed for clinical data and neurodevelopmental testing with the Bayley Scales of Infant Development. A developmental quotient was calculated to compare scores of children at different ages.

RESULTS. Complete data were available on 24 of 50 infants. Surgeries included 14 hemispherectomies and 10 focal resections. Seventeen patients became seizure free; 5 patients had >90% seizure reduction, 1 had >50% seizure reduction, and 1 had no change. The developmental quotient indicated modest postoperative improvement of mental age. The preoperative and postoperative development quotients correlated well. Younger infants had a higher increase in developmental quotient after surgery. Patients with epileptic spasms were younger and had a lower developmental quotient at presentation, but increase in developmental quotient was higher in this subgroup.

CONCLUSIONS. After surgery, seizure frequency and developmental quotient improved. Developmental status before surgery predicted developmental function after surgery. Patients who were operated on at younger age and with epileptic spasms showed the largest increase in developmental quotient after surgery.
Epilepsy surgery is the standard of care for patients with medically intractable focal epilepsy. However, little is known about the impact of epilepsy surgery on development in infants. Decision making about the timing of surgery remains difficult: early surgery may be indicated to prevent recurrent seizures and their devastating effect on development. This, however, implies the risk of unnecessary resection and loss of developing brain tissue in cases where potential medical seizure control may be reached as the brain matures.

Relatively few studies and case reports on surgical and developmental outcome after epilepsy surgery are available in children <3 years old. Two-year postsurgical developmental outcomes were assessed in 24 children with medically intractable infantile spasms who underwent epilepsy surgery. Significant developmental improvement was noted 2 years after surgery. Developmental outcome after surgery was best for patients who received surgery at a younger age and who had the best presurgical developmental scores.

Our aim was to determine the effect of epilepsy surgery in infants (<3 years of age) on development and to describe factors associated with postoperative developmental outcome. The study reports a 12-year experience with infancy epilepsy surgery at the Cleveland Clinic from 1989 to 2001.

METHODS

We identified 50 infants <3 years old among 251 consecutive pediatric patients (<18 years old) undergoing epilepsy surgery at our center between 1989 and 2001. All patients were assessed by a standard protocol including clinical, neuroradiological, neurophysiological, neuropsychiatric, and developmental teams. Each infant was evaluated by using video electroencephalogram monitoring, MRI, and cognitive and developmental assessments. Seizures were classified according to the semiological seizure classification. The data were discussed in a multidisciplinary presurgical meeting. Median follow-up duration after surgery was 6 months. Seizure outcome was assessed by using 4 categories of a modified Engel scale: seizure free, >90% seizure reduction, >50% seizure reduction, and no change in seizure frequency.

Charts were reviewed retrospectively for preoperative and postoperative seizure frequency, neuropsychological testing with the Bayley Scales of Infant Development, antiepileptic drugs (AEDs), surgery type, and pathology. The Bayley Scales of Infant Development (1969) was used with all patients before 1994, at which time the Bayley Scales of Infant Development, second edition, was initiated when it became available at our clinic.

A developmental quotient (DQ; ratio of the Bayley mental age divided by the subject’s biological age × 100) was calculated to compare scores of children at different ages. For example, an infant was tested at 12 months’ biological age. Neuropsychological testing revealed a Bayley mental age of 6 months. The DQ was 6/12 × 100 = 50. Fisher’s exact test, Mann-Whitney \( U \) Wilcoxon test, and Spearman’s correlation coefficient were used for statistics. SPSS 10.0 (SPSS Inc, Chicago, IL) was used to calculate statistics.

RESULTS

Twenty-four patients (18 boys) with complete data and formal neurodevelopmental evaluation were identified among 50 consecutive patients <3 years of age at surgery. Twenty-six patients were excluded because of incomplete neuropsychological data or different preoperative and postoperative neuropsychological tests. Median age at preoperative assessment was 12 months (range: 3.3–33.1 months), median age at surgery was 14 months (range: 3–34 months), and postoperative evaluation was at 24 months (range: 10–53 months), a median of 6 months (range: 4–42 months) after surgery. Median duration between preoperative testing and surgery was 45.5 days (minimum: 1 day; maximum: 408 days) and 195.5 days (minimum: 128 days; maximum: 1259 days) between surgery and postoperative assessment.

Surgeries (13 right, 11 left) included 14 hemispherectomies and 10 focal resections (3 frontal, 3 frontoparietal, 2 parietal, 1 parieto-occipital, and 1 occipital). Pathology consisted of malformation of cortical development (19 patients, 7 with hemimegalencephaly), malformation of cortical development combined with ganglioglioma (2 patients), Sturge-Weber syndrome (2 patients), and tuberous sclerosis (1 patient).

Patients presented with a median of 2 different semio logical seizure types (range: 1–4). Seizure semiology included tonic seizures (15), clonic seizures (15), epileptic spasms (11), eye versive seizures (7), hypomotor seizures (5), and myoclonic seizures (3).

Seizure frequency and number of AEDs decreased after surgery. Before surgery, the patients had a median of 15 seizures per day (range: 0.2–120) and were taking a median of 3 AEDs (range: 0–5). After surgery, seizure frequency decreased to a median of 0 (range: 0–15; \( P < .001 \)), and the number of AEDs was reduced to a median of 1 (range: 1–3; \( P < .001 \)). Seventeen patients became seizure free; 5 patients had >90% seizure reduction (with 0.03–6.00 seizures per day [median: 0.3]), 1 had >50% seizure reduction, and 1 had no change.

Median developmental mental age according to the Bayley scale was 3 months (mean: 5.83) before surgery and 9 months (mean: 11.94) after surgery. The DQ was below average (<100) before and after surgery in all infants. There was a modest postoperative improvement of mental age. It increased from a preoperative median of 37 (range: 0–92) to 49 (range: 2–92) after surgery (\( P < .01 \)). The DQ improved in 17 patients and decreased in 7 patients compared with preoperative assessment (Fig
All 7 infants with no measurable mental development (mental age <1 month on the Bayley scale) before surgery made progress after surgery.

The development of all infants in the study was below that of the average infant (DQ = 100). Only 2 patients were developmentally within the reference range (DQ > 80) before and 3 were within the reference range after surgery. Before surgery, developmental delay (DQ < 70) was present in 22 of 24 children (Table 1). After surgery the number of delayed infants decreased to 18. However, this change was not statistically significant (P = .125, McNemar test). An increase in the number of patients with borderline functioning (DQ = 70–80) accounted for this improvement (0 patients before and 4 after surgery). Profound developmental delay (DQ < 50) was seen in 13 (54%) of 24 infants postoperatively.

Whereas a higher DQ before surgery was correlated with a higher DQ after surgery (correlation coefficient: 0.67; P < .001), infants with a preoperative DQ >50 were more likely to experience a reduction in DQ (6 of 8) than those with a DQ < 50 (1 of 16; P = .01). Although the DQ declined in 7 infants, none of these infants experienced a loss of skills after surgery.

Younger age at the time of surgery was correlated with improvement in the DQ (correlation coefficient: 0.72; P < .001). However, surgery did not affect developmental outcome in infants >12 months of age at time of surgery. The DQ increased after surgery in 10 of the 11 children who had surgery younger than 12 months of age (Fig 2). However, it increased in only 6 of 13 who were older than 12 months of age at the time of surgery (P < .05).

Analysis of seizure semiology revealed that patients with epileptic spasms presented for preoperative assessment and for surgery at a younger age than patients without epileptic spasms. Median age at presentation in patients with epileptic spasms was 4.2 months and in patients without epileptic spasms was 6.1 months (P < .01). Median age at surgery in patients with epileptic spasms was 6.1 months and in patients without epileptic spasms was 19.9 months (P < .01). Preoperative DQ in infants with epileptic spasms was lower (median: 23) than in infants without epileptic spasms (median: 54; P < .01). Change in DQ and improvement was more prominent in the subgroup of infants with epileptic spasms (Fig 3; P < .01). When analyzed separately, only the group of infants with epileptic spasms had a significant improvement in DQ after surgery. There was no difference in the time interval from presurgical assessment to surgery among patients with or without epileptic spasms. Other semiological features were not related to developmental outcome.

Other factors that were not associated with development before or after surgery included preoperative and postoperative seizure frequency, postoperative seizure freedom, and change in number of AEDs, side of surgery, type of resection, or pathology.

CASE ILLUSTRATIONS

Two cases may illustrate the wide spectrum of development in our patient series.

Case 1 (Patient 4)

This patient had seizures since the age of 1 month presenting with right-eye deviation and bilateral eye blinking occurring up to 40 times per day despite treatment with phenobarbital, phenytoin, carbamazepine, and clonazepam. Baseline evaluation at 3.1 months’ biological age revealed a mental Bayley Scales developmental age of 1 month (DQ = 1/3.1 × 100 = 32%). Based on video electroencephalogram, MRI, and fluorodeoxyglucose positron emission tomographic scan data, the patient underwent right parietal resection. Pathology revealed malformation of cortical development. The patient became seizure free and was only maintained on phenobarbital after 6 months. Neuropsychological reassessment at 9.83 months after surgery revealed a developmental age of 9 months (DQ = 92%). The patient caught up, and development at repeat assessment was comparable to a normal infant.

Case 2 (Patient 8)

Left-arm clonic seizures started at the age of 4 months old. The patient continued to have 5 seizures per day despite phenobarbital and carbamazepine. Examination revealed left hemiparesis, left hemianopia, and a port wine stain over the right hemicranium. Baseline neuropsychological assessment at 19.27 months’ biological age revealed a Bayley Scale mental age equivalent of 11 months (DQ = 57%). Based on video electroencephalogram, MRI, and FDG-positron emission tomography.
data, the patient underwent right functional hemispherectomy and became seizure free. Carbamazepine was discontinued. The patient underwent repeat neuropsychological assessment at 27.03 months’ biological age, revealing a Bayley Scales mental age equivalent of 14 months (DQ 52%). Development progressed by 3 months between the first and the second assessment, but not at a normal rate.

DISCUSSION
We present the first study with formal preoperative and postoperative neuropsychological testing limited to infants <36 months of age at the time of epilepsy surgery. All infants (79% with malformation of cortical development) had below average development (DQ < 100) before and after surgery. After surgery, 71% of infants were seizure free, and 92% of patients had at least 90% seizure reduction on reduced AEDs. The median DQ
Seizure frequency and number of AEDs after surgery decreased significantly in all infants. Seventeen of 24 infants became seizure free. This result confirms a study by Chugani et al. that reported seizure freedom or at least 90% seizure frequency reduction in 18 of 23 patients who had focal cortical resections or hemispherectomy between the ages of 5 months to 3.7 years. Our study also reflects previous results from our center that reported 9 of 12 infants with only rare or no epileptic seizures after epilepsy surgery between the ages of 3 and 29 months of age. In addition, the results from our study (71% seizure freedom) also match the seizure freedom rates of extratemporal resections in older children and 67% with seizure freedom or auras in a survey of 2464 patients (adults and children) operated on at 38 different epilepsy centers between 1995 and 1999.

Discontinuation of AEDs after epilepsy surgery has been described in adults and in children, but no studies with focus on children <3 years of age are available. We were able to show that, for the first time, epilepsy surgery in infancy reduces AED treatment. Because seizure frequency reduction and AED reduction both occurred after epilepsy surgery, we are not able to determine whether developmental improvement was related to decreased seizure frequency or AED reduction. It is likely that frequent seizures, as well as the sedating effect of the AEDs, both impair cognitive development. Previous case reports on developmental outcome after epilepsy surgery suggested that early epilepsy surgery in infants with catastrophic epilepsy may allow the resumption of developmental progression during critical stages of brain development and maturation. Mental development tends to progress in the majority of infants after epilepsy surgery, in particular in those with initially no measurable development. In addition, a statistically significant increase occurs in developmental levels at an average age of 21 months after surgery compared with presurgical results. These authors also compared the developmental outcome of their study with all other previously reported infants receiving medical treatment for infantile spasms and found that the developmental outcome in their surgical group was equal and sometimes superior to children treated with either corticotropin or valproic acid. Although overall development remains severely impaired, in our series only a short period of follow-up could be included. Some patients may continue to lose ground and develop at a slower pace, whereas others may actually continue to cross percentiles and continue to catch up. Many infants develop at a faster rate or pick up development but remain abnormal. Meaningful changes may be seen in all infants that develop at a faster rate than their preoperative baseline.

Preoperative and postoperative neurodevelopmental testing in infants <3 years of age undergoing epilepsy surgery is an important tool to predict postsurgical mental outcome and helps to determine the ideal time for resective epilepsy surgery based on presurgical developmental baseline. Developmental outcome was best in children who received epilepsy surgery at a younger age and who had the best baseline assessments before epilepsy surgery. One hypothesis that explains the better presurgical test results in children who undergo surgery when they are <1 year of age is that there is less time for seizures to influence development. Early treatment and seizure control seem to be key to improved developmental outcome. This has also been confirmed by a recent study on long-term cognitive outcomes of a cohort of children with infantile spasms of unknown etiology that were treated with high-dose corticotropin. Twenty-two infants were treated within 1 month of onset of infantile spasms and 15 after 1 to 6 months. All 22 infants of the early treatment group, but only 40% in the late treatment group, had normal cognitive outcome. In addition, infants with only minimal mental retardation at presentation were more likely to have a normal cognitive outcome between the ages of 6 to 21 years of age.

Seizure semiology may correlate with outcome after epilepsy surgery. Recent work on the developmental outcomes after epilepsy surgery that also included older children at the time of surgery suggested that postoperative DQ calculated based on the Vineland Adaptive Behavior Scale was better in patients with epilepsy of shorter duration and earlier surgical intervention. In this series, infants were classified based on medically refractory spasms, successfully treated spasms, and no epileptic spasms. Interestingly, age at surgery was older with fewer documented spasms. This correlates with our results that showed more spasms in the patients that underwent early operative intervention. Epileptic spasms present earlier and, therefore, may just be an indicator for severe developmental delay at the time of presentation and early recognition and treatment of epilepsy. It remains unclear whether epileptic spasms are a separate risk factor or whether the analysis of patients with spasms may be confounded by the early time of presentation and surgery.

Pathologic findings in our group may have influenced our data. The most prominent pathologic diagnosis in our group consisted of malformation of cortical development, which was shown to predispose for less improvement in language and intelligence scores on follow-up after hemispherectomy in a series of older children (age at hemispherectomy ranged from 4 months to 20 years). Two patients with Sturge-Weber syndrome did not improve developmentally after epilepsy surgery de-
spite seizure freedom in our series. A previous case series describes the IQ after hemispherectomy in a case series of patients with Sturge-Weber syndrome on development. Five of 6 patients had a favorable outcome with intelligence quotients >80%, with follow-up ranging from 1 to 13 years. The sixth patient was the only patient in this series who did not undergo hemispherectomy during the first year of life, and this patient was the only one that had no improvement in his IQ. These results correspond to our experience that children undergoing epilepsy surgery during the first year of life may have a better developmental outcome (Fig 2).

A limitation of our series is the lack of an appropriate control group. To compare presurgical and postsurgical development in our series, we had to use the hypothetical construct of the DQ to compare presurgical and postsurgical development. The retrospective study approach led to variable intervals between preoperative and postoperative testing, selection bias because of data acquisition at a tertiary epilepsy center with referral bias and ascertainment bias, as well as limitations of the applied test scales. In addition, neuropsychological data were incomplete in half of the initially included 50 cases. Although our patients were highly selected, it seems unlikely that they were selected to have favorable outcome, because infants were only selected for surgical intervention after they had failed several antiepileptic medications.

CONCLUSIONS
We present the first study, to our knowledge, with formal preoperative and postoperative neurodevelopmental testing limited to infants <36 months old at the time of epilepsy surgery. The median DQ improved after surgery for the group overall, and individual DQs improved for 71% of infants. Developmental status before surgery predicted developmental function after surgery. Assessment of DQ change in relation to age at surgery and seizure semiology will require a refined prospective study design comparing surgical and medical treatment of seizures in infants.

REFERENCES
ABSTRACT

OBJECTIVE. Our goal was to determine the value of measuring plasma caffeine levels in preterm neonates treated with caffeine for apnea. We evaluated plasma concentrations of caffeine attained in preterm neonates at standard doses, at varying postconceptual ages, with renal or hepatic dysfunction and when there was clinical lack of efficacy. We hypothesized that measurement of plasma caffeine concentrations during apnea therapy is not clinically helpful.

PATIENTS/METHODS. An observational study was conducted at Hutzel Women’s Hospital between January 2000 and September 2005. Preterm neonates who were being treated with caffeine and who had a plasma caffeine level measured on at least 1 occasion were included.

RESULTS. A total of 231 caffeine blood levels were obtained from 101 preterm neonates with a median gestation of 28 weeks (range: 23–32 weeks) and birth weight of 1030 g (range: 540–2150 g). The caffeine citrate dose used ranged form 2.5 to 10.9 mg/kg (median: 5 mg/kg), and the levels ranged from 3.0 to 23.8 mg/L. Levels were between 5.1 and 20 mg/L in 94.8%, <5 mg/L in 2.1%, and >20 mg/L in 3.1%. Levels in the 5.1 to 20 mg/L range were attained on 91.3% of occasions when there was concomitant renal dysfunction (n = 23) and in all cases of hepatic dysfunction (n = 13). The median (25th, 75th quartiles) levels drawn for lack of efficacy (14.1 [10.2, 8.3] mg/L; n = 94) were comparable to those obtained for routine monitoring (13.7 [11, 9] mg/L; n = 107).

CONCLUSIONS. A majority of preterm neonates attain plasma caffeine levels between 5 and 20 mg/L, independent of gestation. This observation held even for the small number of subjects with elevated blood urea nitrogen, serum creatinine, or liver enzyme levels. Therapeutic drug monitoring is not necessary when caffeine is used for the treatment of apnea of prematurity in neonates.
CAFFEINE is a methylxanthine used extensively as first-line pharmacotherapy in apnea of prematurity. Its efficacy in reducing the frequency of apneic episodes and the need for mechanical ventilation is well established.\(^1\)\(^2\) More recently, a large randomized, controlled trial among very low birth weight infants demonstrated a significant decrease in the rate of bronchopulmonary dysplasia in the caffeine-treated group, with a low incidence of short-term toxicity.\(^3\)

The recommended standard dosing for caffeine citrate is 20 to 40 mg/kg loading followed by 5 to 8 mg/kg per day as maintenance.\(^4\) Larger doses up to 20 mg/kg day in the periestubation period have shown higher rates of successful extubation, without evidence of harm in the first year of life.\(^5\) Pharmacokinetic studies in premature neonates have established that the half-life of caffeine is prolonged to 102.9 ± 17.9 hours and remains prolonged for as long as 38 weeks' gestation, which reflects the maturational deficit of its hepatic biotransformation in the newborn.\(^6\) The transition to adult levels of elimination occurs at ~3 to 4½ months.\(^7\) Other factors such as cholestatic jaundice and breastfeeding seem to further prolong the half-life of caffeine.\(^8\) Clinically effective plasma concentrations vary over a wide range of 5 to 50 mg/L and overlap with subtherapeutic concentrations.\(^8\)\(^9\)

Despite extensive pharmacokinetic data, the value of measuring caffeine plasma concentrations in neonates is unclear. Monitoring of serum concentrations of caffeine if there is lack of clinical response or if there is suspected toxicity has been recommended.\(^10\) Many centers also perform routine drug monitoring every week or 2 weeks, especially in extremely premature infants.\(^8\)\(^11\)\(^12\) We undertook an observational study to determine the value of measuring plasma caffeine levels in preterm neonates being treated with caffeine for apnea. Our specific aim was to evaluate the range of plasma caffeine levels attained in preterm neonates at varying postnatal and postconceptional ages (PCAs) or with renal or hepatic dysfunction, using standard maintenance doses. We also evaluated the plasma drug concentrations obtained in the subgroup of subjects in whom the indication for determining levels was lack of clinical efficacy. Because of the wide therapeutic index of caffeine and the dose-plasma concentration relationship in newborn infants with apnea, we hypothesized that measurement of plasma concentrations of caffeine during apnea therapy is not clinically helpful.

PATIENTS AND METHODS

This was an observational study in which preterm infants treated with caffeine for apnea of prematurity, who had ≥1 blood levels obtained between January 2000 and September 2005 were identified. Clinical data, as well as plasma caffeine concentrations, were then included in the pharmacy database at Hutzel Women’s Hospital, and the data were collated and analyzed. The institutional review board at Wayne State University approved the protocol. Data on the gestational age, birth weight, dose, and age at initiation of caffeine therapy; age at the time of obtaining the blood level; and renal and liver function profile within 3 days of obtaining the level were abstracted from the charts. The clinical indications for the drug level were noted from the medical chart when documented and classified as 1 of the following: lack of efficacy, suspected toxicity, or routine monitoring (trough level obtained on day 5 of therapy). The duration of caffeine therapy was noted.

The decisions on when and whether to start caffeine, the doses, and need for and timing of plasma levels were taken by the attending neonatologist. Blood samples for assay of caffeine levels were usually drawn 2 hours before the next dose. Plasma caffeine levels were measured by the enzyme immunoassay technique (Dade Behring Inc, Deerfield, IL), and the results are reported in milligrams per liter.

SPSS 14 (SPSS, Chicago, IL) was used to obtain the median and quartile ranges of caffeine blood levels at different doses and PCAs.

RESULTS

A total of 137 infants were treated with caffeine during the study period. Thirty-six of these were excluded because therapeutic drug monitoring was not performed, either because the apnea was well controlled at standard doses, the infant expired or was transferred out of the institution, or because of the clinical neonatologist’s preference. Data from a total of 101 neonates were abstracted, on whom 231 caffeine levels were obtained.

Clinical Profile

The median gestational age at birth was 28 weeks (range: 23–32 weeks), and the median birth weight was 1030 g (range: 540–2150 g). Caffeine was initiated at a median age of 6 days (range: 1–70 days), with the third day of life being the most frequent day of initiation of therapy. The median dose of caffeine citrate was 5.0 mg/kg (range: 2.5–10.9 mg/kg). Caffeine levels were performed at a median PCA of 31 weeks (range: 24–41 weeks). The initial level was drawn at a median postnatal age of 14 days (range: 4–79 days) and a median PCA of 30 weeks (range: 24–39 weeks). Plasma drug concentrations of caffeine were obtained at a PCA of <28 weeks in 32 sample (13.8%), between 28 and 32 weeks in 132 (56.9%), and beyond 32 weeks in 67 (29.3%) of samples. The median number of levels per patient was 2 (range: 1–8). Of the 101 infants, 45 had a single level, 17 had 2 levels, 16 had 3 levels, 15 had 4 levels, 5 had 5 levels, and 1 infant each had 6, 7, and 8 levels obtained. The median duration of caffeine therapy was 29 days (range: 6–75 days).
Caffeine Plasma Drug Concentrations
The median caffeine concentration was 10.7 mg/L (SD: 4.0; range: 3–23.8 mg/L). The 25th quartile was 8.7 mg/L, and the 75th quartile was 14 mg/L. Figure 1 shows the distribution of levels in the study population and at different PCAs.

Effect of Renal and Hepatic Function on Caffeine Drug Concentrations
On the 23 occasions when serum creatinine was ≥1 mg/dL and/or blood urea nitrogen (BUN) was >30 mg/dL, the mean caffeine level was 13.9 (4.1) compared with 11.5 (4.1) mg/L on the 147 occasions when they were not. Plasma caffeine concentrations were within the 5 to 20 mg/L range on all 13 occasions when aspartate aminotransferase or alanine aminotransferase was between 9, 13.7, and >22.7 mg/L. The highest BUN and creatinine in our data set was 78 mg/dL, and the 75th quartile was 14 mg/L. Figure 1 depicts multiple levels at different PCAs, normalized to a dose of 1 mg/kg in the infants who had ≥3 levels.

Indications for Caffeine Therapeutic Drug Monitoring
The indications for obtaining caffeine levels were documented on 203 occasions. They were performed for lack of efficacy in 94 (46.3%) cases, routine monitoring in 107 (52.7%) cases, and for tachycardia (adverse effect) on 2 (1%) occasions. The median (25th, 75th quartiles) caffeine levels when performed for lack of efficacy, routine monitoring were 10.2 (8.3, 14.1) and 11 mg/L (9, 13.7), respectively. The levels on the 2 occasions with tachycardia were 13.9 and 22.7 mg/L.

Effect of Dose, Age, and Maturity on Caffeine Plasma Concentrations
The mean (SD) caffeine doses in the subgroups with plasma concentrations <5 mg/L, between 5.1 and 20 mg/L, and >20 mg/L were comparable at 5 (1.6), 5.2 (0.9), and 5.2 (1.1) mg/kg. The mean level obtained at a maintenance dose of 2.5 to 5 mg/kg was 11.18 mg/L (SD: 4.02 mg/L; n = 123), between 5.1 and 7.5 mg/kg was 12 mg/L (3.97 mg/L; n = 97), and >7.5 mg/kg was 9.86 mg/L (SD: 5.14 mg/L; n = 7). Table 2 shows the median and ranges of PCAs, caffeine dose, and multiple plasma levels. The number of infants who had ≥5 or levels were very few, which restricts the application of a repeated measures analysis of variance model to the smallest sample size in a cell level. For example, if we were to include level 5 data across the other previous levels, the sample size comparison would be restricted to a similar sample, that being an n of 8. We did, however, conduct the analysis by choosing a general linear model and found no statistically significant differences. Figure 2 depicts multiple levels at different PCAs, normalized to a dose of 1 mg/kg in the infants who had ≥3 levels.

DISCUSSION
Our data suggest that the overwhelming majority (95%) of preterm neonates achieve plasma concentrations of caffeine between 5 and 20 mg/L at standard doses, independent of low gestation. This observation held even for the small number of infants who had elevated liver enzyme, serum creatinine, and/or BUN levels. In the subgroup of infants in whom caffeine plasma concentrations were obtained for lack of clinical efficacy, three quarters of the levels were <15 mg/L, which suggests that higher doses and plasma concentrations may be required for optimal efficacy.

Caffeine plasma concentrations achieved by the pre-term neonates in our study were in accord with published pharmacokinetic data. Plasma concentrations between 7.4 and 19.4 mg/L (mean level: 13.7 mg/L) were reported at a maintenance dose (caffeine base) of 2.5 mg/kg in an early dose-finding study. A large population pharmacokinetic study reported mean serum caffeine concentrations of 7.4, 35.8, and 69 mg/L in 119 preterm neonates who were given 3, 15, and 30 mg/kg, respectively, once daily for a week. Plasma concentrations among specific ethnic groups (range: 3.6–28.4 mg/L with means between 10 and 20 mg/L) also fall within the same range. A previous population pharmacokinetic study based on routine therapeutic drug monitoring in 60 neonates reported that plasma concentrations within the traditional target range (5–20 mg/L) should be achieved in >70% of neonates. Another study on 16 preterm neonates found plasma levels within the 13 to 20 mg/L range in 69% of cases. Our study results were derived from a larger patient population in whom levels were obtained by the clinician for
lack of efficacy, concerns about delayed clearance because of hepatic or renal dysfunction or extreme prematurity and rarely, adverse effects. It is reassuring that even in these specific situations, plasma concentrations of caffeine were as expected, and indeed within the recommended therapeutic range in 90% of infants.

This study also attempts to address the implication of a specific plasma level of caffeine. There is uncertainty on the precise desired plasma concentration and its correlation with efficacy. Although a decrease in apnea and increase in respiratory drive is known at plasma concentrations as low as 2.9 and 4 mg/L, optimal effect is at 10 mg/L. Higher doses and caffeine levels have been targeted in a few previous studies with some benefit and no adverse effects. A high dose regimen of 25 mg/kg loading followed by 6 mg/kg daily maintenance caffeine base generated mean (SD) plasma concentrations of 30.4 (4.0) mg/L and a more rapid reduction in apneas within 8 hours.16 Lee and associates,13 in a population pharmacokinetic study used daily caffeine citrate doses as high as 15 and 30 mg/kg over 7 days and attained mean caffeine levels of 35.8 and 69 mg/L, with no untoward effects. They suggested that preterm infants may require serum concentrations of 35 mg/L for effective prophylaxis against apnea after extubation. Our data set consisted of a large number of levels (n = 94), which had been obtained for a clinical concern of lack of efficacy. The median level in this subgroup was 10.2 mg/L, with a 75th quartile of 14.1 mg/L. These data taken collectively suggest that there are some preterm neonates who require daily doses of caffeine greater than previously recommended and that higher plasma levels are tolerated well in the short-term, although there are no long-term studies.

There was no correlation between caffeine levels and dose in our study within the relatively narrow dose range (5–8 mg/kg) used, probably because of the confounding effects of age and maturity. This relationship has been well established by other investigators, however.13,17 Plasma levels remained stable at later postnatal ages in infants in whom multiple levels were obtained. Because the PCA in most of our study group was 38 weeks, this was as expected.8

The limitations of our study are that the numbers of subjects who had renal or hepatic dysfunction at the time of a caffeine level were small. We were not able to evaluate efficacy in a systematic manner because the number of apneic episodes were not precisely recorded in the chart.

CONCLUSIONS

Routine monitoring of plasma concentrations of caffeine is unnecessary, even in extremely premature infants with renal or hepatic dysfunction or after prolonged use, because the overwhelming majority achieve concentrations in the range of 5 to 20 mg/L. In the subgroup of infants who do not show a clinical response to standard doses of caffeine, higher plasma levels may be targeted. Monitoring of plasma levels may be prudent in these cases. The risk–benefit of this approach needs to be studied in additional prospective clinical trials. Elimina-

### TABLE 1 Distribution of Caffeine Levels in the Presence of Renal or Hepatic Dysfunction

<table>
<thead>
<tr>
<th>Plasma Level, mg/L</th>
<th>BUN &gt;30 mg/dL or Serum Creatinine &gt;1 mg/dL</th>
<th>Normal Renal Function</th>
<th>AST/ALT &gt;60 or Direct Bilirubin &gt;1 mg/dL</th>
<th>Normal Hepatic Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;5</td>
<td>—</td>
<td>5 (3.4)</td>
<td>—</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td>5.1–20</td>
<td>21 (91.3)</td>
<td>138 (93.9)</td>
<td>13 (100)</td>
<td>85 (95.5)</td>
</tr>
<tr>
<td>&gt;20</td>
<td>2 (8.7)</td>
<td>4 (2.7)</td>
<td>—</td>
<td>3 (3.4)</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>147</td>
<td>13</td>
<td>89</td>
</tr>
</tbody>
</table>

AST indicates aspartate aminotransferase, ALT, alanine aminotransferase, —, 0 patients. Values represent the number (percentage) of samples.

### TABLE 2 Comparison of Multiple Plasma Levels

<table>
<thead>
<tr>
<th>Sample</th>
<th>1 (n = 100)</th>
<th>2 (n = 55)</th>
<th>3 (n = 39)</th>
<th>4 (n = 23)</th>
<th>5 (n = 8)</th>
<th>6 (n = 3)</th>
<th>7 (n = 2)</th>
<th>8 (n = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level, mg/L</td>
<td>10.8 (3–19.9)</td>
<td>11 (3.5–23.4)</td>
<td>10.8 (4.1–23.8)</td>
<td>10.6 (4.9–23.5)</td>
<td>9.0 (5.7–12.8)</td>
<td>12 (9.8–13.9)</td>
<td>15.8 (10.9–20.8)</td>
<td>9.8</td>
</tr>
<tr>
<td>Dose, mg/kg</td>
<td>4.9 (3.6–9.4)</td>
<td>5 (2.5–7.3)</td>
<td>5 (3.2–8.0)</td>
<td>5.1 (3.5–8.8)</td>
<td>5.2 (4.0–6.0)</td>
<td>6.6 (5.5–7.2)</td>
<td>8.4 (6.6–10.1)</td>
<td>6.7</td>
</tr>
</tbody>
</table>

Values shown are median (range). The values in parentheses represent the number of blood samples for caffeine plasma concentrations.

![Figure 2](http://example.com/fig2.png)

**FIGURE 2**

Trends of plasma caffeine concentrations at different PCAs in infants who had ≥3 levels, normalized to a dose of 1 mg/kg.
tion of the practice of obtaining plasma levels would result in less blood draws and considerable cost-savings at a cost of $50 per assay.

REFERENCES

HOPES SOAR AFTER RECORD HOSPITAL GIFT OF $400 MILLION

“South Dakota’s economic renaissance has gone largely unnoticed. . . . T. Denny Sanford, a low-key billionaire who made his home and fortune here, will help sustain the state’s economic boom with a $400 million gift to the Sioux Valley Hospitals and Health System, the state’s largest employer. Hospital officials hope the gift—the largest ever to a hospital, according to the Center for Philanthropy at Indiana University—will help transform Sioux Valley Hospitals, which will change its name to Sanford Health, into a national institution that will eclipse Johns Hopkins and the Mayo Clinic. ‘He told me he doesn’t want this to be just another Mayo,’ said Kelby K. Krabbenhof, Sioux Valley’s chief executive. It has four stated goals: to build five pediatric clinics around the country; to expand research, especially in pediatrics; to build a health care campus with more than 20 separate facilities; and to identify a promising line of medical research and follow it to a cure, much the same way John D. Rockefeller’s money found a cure for yellow fever and Bill Gates is searching for a cure for HIV/AIDS.”

Noted by JFL, MD
Maternal Antibodies in Breast Milk Protect the Child From Enterovirus Infections

Karita Sadeharju, MD, PhD, a Mikael Knip, MD, PhDb,c, Suvi M. Virtanen, MD, PhDd, Erkki Savilahti, MD, PhD, e Sisko Tauriainen, PhD, a Pentti Koskela, MD, PhD, a, Hans K. Åkerblom, MD, PhD, d Heikki Hyöty, MD, PhD, f and the Finnish TRIGR Study Group

aDepartment of Virology, University of Tampere and Tampere University Hospital, Tampere, Finland; bHospital for Children and Adolescents, University of Helsinki, Helsinki, Finland; cDepartment of Pediatrics, Tampere University Hospital, Tampere, Finland; dDepartment of Epidemiology and Health Promotion, National Institute of Public Health, Helsinki, and School of Public Health, University of Tampere, Tampere, Finland; eNational Institute of Public Health, Oulu, Finland

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

ABSTRACT

OBJECTIVE. Enterovirus infections are frequent in infants and may cause severe complications. We set out to assess whether breastfeeding can protect against these infections and whether such an effect is related to maternal antibodies in breast milk or in the peripheral circulation of the infant.

METHODS. One hundred fifty infants who were prospectively followed up from birth were monitored for enterovirus infections. The duration of breastfeeding was recorded, and maternal breast milk and blood samples were regularly taken at 3-month intervals for the detection of enterovirus antibodies and RNA. Maternal serum was available from early pregnancy, delivery, and 3 months postpartum.

RESULTS. Enterovirus infections were frequent and were diagnosed in 43% of infants before the age of 1 year and in 15% of the mothers during pregnancy. Infants exclusively breastfed for >2 weeks had fewer enterovirus infections by the age of 1 year compared with those exclusively breastfed for ≤2 weeks (0.38 vs 0.59 infections per child). High maternal antibody levels in serum and in breast milk were associated with a reduced frequency of infections. This effect was seen only in those infants breastfed >2 weeks, indicating that breast milk antibodies mediate this effect. Enterovirus RNA was not found in any of the breast milk samples.

CONCLUSIONS. These results suggest that breastfeeding has a protective effect against enterovirus infections in infancy. This effect seems to be mediated primarily by maternal antibodies in breast milk.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-0780
doi:10.1542/peds.2006-0780

Key Words enterovirus infections, maternal antibodies, autoantibodies, breastfeeding

Abbreviations

RT-PCR—reverse transcriptase polymerase chain reaction
IgG—immunoglobulin G
IgA—immunoglobulin A
CBV4—coxsackievirus B4
EV11—echovirus 11
EIA—enzyme immunoassay
EIU—enzyme immunoassay unit

Accepted for publication Dec 22, 2006
Address correspondence to Heikki Hyöty, MD, PhD, Department of Virology, Medical School, University of Tampere, Biokatu 10, 33520 Tampere, Finland. E-mail: heikki.hyoty@uta.fi

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics

941
Earlier studies have suggested that breastfeeding may protect the infant against infections. Breast milk contains several compounds that may have antimicrobial effects. Maternal antibodies in breast milk were proposed to play a major role, but other factors such as lactoferrin may also be important. In addition to antibodies in breast milk, systemic transplacentally transported antibodies contribute to the protection of the child.

Enterovirus infections are frequent in childhood. Most infections are mild or even asymptomatic, but complications are also quite common including meningitis, myocarditis, otitis media, and severe systemic infections in newborn infants. In addition, enteroviruses were suspected of playing a role in certain chronic diseases, such as type 1 diabetes and cardiomyopathies. Altogether, because of their frequent occurrence and relatively severe complications, enterovirus infections have considerable clinical relevance.

In our study, we wanted to assess whether breastfeeding can protect infants against nonpolio enterovirus infections. Although this question is of unequivocal clinical importance, only 1 study has so far been published. That study found an epidemiologic association between lack of breastfeeding and frequent enterovirus infections in young infants (<1 month of age), but it did not specifically assess the role of breast milk antibodies in this phenomenon. Accordingly, the possible protection conferred by breastfeeding needs to be confirmed. In addition, it is not known whether the protection is mediated by maternal antibodies in breast milk, other breast milk compounds, or transplacentally transferred antibodies. It is also not known how common is the transmission of enteroviruses from the mother to the child through breast milk, because only 1 study was published on this subject on the basis of a small series of children. The transmission through breast milk of some viruses like HIV and cytomegaloviruses was demonstrated.

We were able to address these issues in a unique prospective series of infants who were observed sequentially starting from birth. Details about breastfeeding were recorded by questionnaires during the follow-up, and serial blood samples were taken from the infant to diagnose enterovirus infections by using a wide panel of assays (reverse transcriptase polymerase chain reaction [RT-PCR] and antibody measurements). In addition, cord blood and breast milk samples were available for the measurement of maternal enterovirus antibodies and direct detection of viral RNA.

Subjects and Methods

Study Population

The study population was comprised of 150 infants and their mothers from families with at least 1 member affected by type 1 diabetes. These infants carried increased HLA-conferred susceptibility to type 1 diabetes (HLA-DQB1 *0302 and/or *02 but no protective allele). The families took part in the second pilot study of the TRIGR (Trial to Reduce Insulin-Dependent Diabetes Mellitus in the Genetically at Risk) Project in Finland. This randomized, double-blind trial was intended to evaluate the feasibility of a nutritional intervention trial in infancy and to observe the possible effect of weaning to a highly hydrolyzed formula on the emergence of diabetes-associated autoantibodies. The pilot protocol was previously described in detail.

The infants were born between March 1995 and December 1996 in 15 hospitals around Finland. In addition to the cord blood sample, blood samples were obtained from children at follow-up visits at the ages of 3, 6, 9, 12, 18, 24, and 36 months. Clinical symptoms were not recorded, but the children were healthy at the visits. More than 80% of the children (n = 125; 83%) were observed until the age of 24 months and 51% (n = 77) until the age of 36 months (average observation period: 27.6 months; range: 3–36 months). Eighteen children (12%) developed signs of β-cell autoimmunity (≥1 type 1 diabetes-associated autoantibodies), and 8 had progressed to clinical type 1 diabetes by the age of 36 months.

Maternal blood samples were obtained at the end of the first trimester of pregnancy, at the time of delivery, and at 3 months after delivery. Breast milk samples were taken after delivery (mean: 2 days’ postpartum; range: 0–3 days) and 3 months later (range: 1.9–4.0 months). Only 1 breast milk sample was available from 66 mothers, and both samples were available from 84 (56%) mothers.

All samples were stored at −20°C until analyzed. Written informed consent was obtained from the mothers before enrollment. The study was approved by the Joint Ethics Committees of the participating hospitals.

Virus Analyses

Enterovirus Antibodies

Immunoglobulin G and immunoglobulin A (IgG and IgA) class antibodies were measured against purified coxsackievirus B4 (CBV4), purified echovirus 11 (EV11), and a synthetic enterovirus peptide antigen (sequence KEVPALTAVETGAT-C derived from an immunodominant region of capsid protein VP1, which is a common epitope for several enteroviruses) by using enzyme immunoassay (EIA) as described earlier.

The sensitivity and specificity of these antibody assays were validated in earlier studies by analyzing virus-isolation of neutralizing antibody-confirmed infections, indicating that these assays can detect 50% to 70% of enterovirus infections with close to 100% specificity. The purified CBV4 and EV11 were incubated at 56°C for 15 minutes to expose antigenic determinants common
for various enterovirus serotypes. Serum and breast milk samples were analyzed at 1/100 (IgA), 1/500 (EV11-IgG), and 1/2000 (other IgG assays) dilutions in phosphate-buffered saline supplemented with 1% bovine serum albumin and 0.05% Tween 20. The results of antibody measurements (OD490) were expressed in enzyme immune units (EIUs), showing the relative antibody reactivity of the sample in relation to the positive and negative reference sera included in each assay.

**Enterovirus RNA**
RNA was extracted by a QIAamp viral RNA kit (Qiagen, Hilden, Germany) from 140 µL of serum or breast milk that was not previously thawed, according to the manufacturer’s protocol. Enterovirus RNA was detected by repeated RT-PCR and a subsequent hybridization, which detects practically all enterovirus serotypes, as previously described. All positive samples were confirmed by repeated RT-PCR.

**Diagnostic Criteria for an Enterovirus Infection**
A twofold or greater increase in the antibody level against an antigen observed between 2 consecutive serum samples from infants or mothers and exceeding the cutoff level of seropositivity (15 EIUs) was considered an infection according to the criteria validated in earlier studies. Because the sera were analyzed by using a single dilution (and not end-point titration), a twofold increase represented quite a noticeable change in antibody levels. The presence of enterovirus RNA in serum was taken as a marker of a current infection.

**Statistical Analyses**
The ages at the time of infection, frequency of infections, and enterovirus antibody levels were compared by using the Mann-Whitney U test. Correlations were analyzed by using the nonparametric Spearman’s correlation test. P values < .05 were considered statistically significant.

**RESULTS**

**Frequency of Enterovirus Infections**
An enterovirus infection was diagnosed during pregnancy in 23 mothers (15%) and during the first 3 months after delivery in 8 mothers (5%). A total of 149 separate enterovirus infections were diagnosed in the children. Seventy-three children (49%) had 1 infection, whereas 35 children (23%) had 2, and 2 children (1%) had 3 infections. Forty-three percent (64 of 150) of the infants had at least 1 infection before the age of 1 year. Twelve percent (18 of 150) of all children were enterovirus RNA-positive in serum, and 3 of them had >1 RNA-positive sample. A rise in enterovirus antibody levels coincided with the detection of viral RNA in serum on 8 occasions (44%), and 5 additional RNA-positive samples (28%) were taken before or at the age of 6 months, when maternal antibodies are able to mask antibody rises. As an example, Fig 1 shows results from the virus analyses of 1 child and his/her mother.

**Breastfeeding and Enterovirus Infections**
Infants were exclusively breastfed for an average of 2.6 months (range: 0–6 months), and 18% of the children were exclusively breastfed for > 4 months. The average duration of the whole breastfeeding period (exclusive and nonexclusive) was 8.1 months (range: 0.15–23 months). Seventy percent of the children had breast milk for >6 months and 12% for >1 year.

The children exclusively breastfed for 2 weeks or less (N = 38) tended to have more enterovirus infections compared with the children with a longer duration of exclusive breastfeeding (N = 112). This difference was seen in infections diagnosed before the age of 1 year (0.59 vs 0.38 infections per child; P = .04) but not in infections diagnosed at an older age (0.54 vs 0.49 infections per child, not significant). The average duration of exclusive breastfeeding was shorter in those infants who experienced at least 1 enterovirus infection before the age of 1 year compared with the infants who had no infections during their first year (median: 2.1 vs 2.9 months, respectively; P = .05). However, the duration of the total breastfeeding period did not differ between the 2 groups (median: 7.6 vs 8.5 months, not significant). The infants who had enterovirus RNA detectable in their serum before the age of 1 year (N = 18) also had a shorter exclusive breastfeeding period than the other infants (median duration: 1.0 vs 3.0 months; P = .04).
Maternal Antibodies and Enterovirus Infections

Maternal IgG class enterovirus antibody levels were lower in serum at delivery compared with early pregnancy or 3 months after delivery (Fig 2). A strong correlation was seen between IgG enterovirus antibody levels in cord blood and in maternal serum (CBV4-IgG: $r = 0.94$; enterovirus peptide IgG: $r = 0.90$; EV11-IgG: $r = 0.86$). The offspring of mothers who had particularly high CBV4 antibody levels (exceeding the 75th percentile) at delivery had fewer enterovirus infections before the age of 1 year compared with the offspring of mothers with lower antibody levels (0.32 vs 0.56 infections per child; $P = .01$, Fig 3). The same phenomenon was observed when antibodies against the common enterovirus peptide antigen were used as an indicator of enterovirus immunity (0.38 vs 0.60 infections per child, respectively; $P = .001$).

Infants whose mothers had high enterovirus IgA levels (over the 75th percentile) in breast milk at delivery had fewer enterovirus infections by the age of 1 year (0.33 vs 0.67 infections per child; $P = .001$; Fig 3). As expected, breast milk samples taken a few days after delivery (colostrum) had higher IgA class enterovirus antibody levels than samples taken ~3 months later (median: 38 vs 12 EIUs; $P = .0001$; Fig 4). However, the antibody levels increased exceptionally by >50% in 5 mothers, which may reflect a recent enterovirus infection. In fact, an enterovirus infection was diagnosed during pregnancy in 2 of them. This phenomenon was seen when the mothers who had an enterovirus infection during pregnancy were compared with the other mothers: the former group had higher IgA class enterovirus antibody levels in breast milk (median IgA: 76 vs 38 EIUs; $P = .005$).

Because a clear correlation was seen between enterovirus antibody levels in breast milk and maternal serum, we wanted to analyze which maternal antibodies are more important for the protection of the offspring, those transferred by breast milk or those transplacentally transferred. We analyzed the effect of these antibodies in 2 groups. One group ($N = 38$) included infants who were exclusively breastfed for ~2 weeks, which would not provide the child with significant protection by breast milk (discussed earlier in this article), whereas the other group ($N = 112$) included children who were breastfed for a longer period. In the group with a short duration of breastfeeding, the levels of enterovirus antibodies in the maternal serum and breast milk were not related to the frequency of enterovirus infections among these infants. However, in the group with a longer duration of breastfeeding, the mothers whose offspring experienced an enterovirus infection by the age of 1 year had lower maternal enterovirus antibody levels both in serum and in breast milk at delivery than the mothers whose offspring had no signs of an infection by that age (medians in serum CBV4 IgG: 30 vs 50 EIUs; $P = .007$ and in breast milk IgA: 29 vs 38 EIUs; $P = .01$).

Enterovirus RNA in Breast Milk

None of the breast milk samples ($N = 234$) was positive for enterovirus RNA. Eight mothers were enterovirus RNA positive in serum at the same time as the breast milk sample was taken.
Maternal immunity is transferred to the infant by antibodies. Both transplacentally transported maternal IgG, as well as breast milk IgA, can protect the child against infections by neutralizing the infectivity of microbes (Fig 3). In addition, breast milk contains other antimicrobial substances, which may also play a role. Breastfeeding decreases morbidity in gastroenteritis, septicemia, otitis media, celiac disease, and other diseases. Despite this evidence, surprisingly few studies are available with the effect of breastfeeding evaluated prospectively by using laboratory-confirmed microbial diagnosis. In addition, the relative significance of breast milk antibodies and transplacentally transferred antibodies remains unknown.

Enteroviruses are important pathogens, being one of the most frequent causes of infections in children. Despite their considerable clinical significance, there are no protective treatments available except for polioviruses, and, therefore, it is important to assess the possible beneficial effect of breastfeeding in this respect. Only 1 study has addressed this question. In that study, children were followed up for a much shorter period than in our study (from birth until the age of 1 month). Enterovirus infections were diagnosed by virus isolation from serial stool samples whereas serum samples were not analyzed. As in our study, these investigators found that the absence of breastfeeding was associated with an increased risk of enterovirus infections. We had a unique opportunity to assess the possible protective effect in relation to maternal antibody levels in serum and breast milk. The protection conferred by breastfeeding was related to the levels of IgA class enterovirus antibodies in breast milk, which seemed to be more important than transplacentally acquired antibodies. This is logical, because the primary replication of enteroviruses occurs in gut mucosa, where breast milk antibodies can directly neutralize the virus in the intestine and prevent its subsequent spread to the circulation. The protective effect was related to exclusive breastfeeding rather than total breastfeeding, suggesting that there may be a dose-effect related to the greater amount of breast milk ingested by exclusively breastfed infants. In addition, protection correlated with enterovirus IgA levels, supporting true biological effect rather than the influence of potential confounding factors, like family size or socioeconomic status. Eighteen children developed signs of β-cell autoimmunity or clinical type 1 diabetes during the follow-up, but the protective effect of maternal antibodies was seen irrespective of the diabetes status. However, we observed previously that these children experienced more enterovirus infections compared with other children of this cohort, suggesting that enterovirus infections may increase the risk of β-cell autoimmunity.

Enterovirus antibody levels were on average >2 times higher in breast milk samples taken a few days after delivery than 3 months later (Fig 4). This is in line with the fact that the initial breast milk, colostrum, has particularly high concentrations of antibodies, which then gradually decrease over the first 2 weeks’ postpartum and remain fairly stable throughout subsequent lactation. Because of the high levels of neutralizing antibodies in colostrum, the first dose of oral poliovirus vaccine is not given just after birth, because these antibodies could decrease the efficacy of the live vaccine. After the age of 6 weeks, breastfeeding no longer has such an inhibitory effect on the oral poliovirus vaccine.

Some viruses, such as HIV and cytomegalovirus, can be transmitted via breast milk. In our study, enterovirus RNA was not detected in any of the breast milk samples, although many mothers had an infection during lactation and some of them were even positive for enterovirus RNA in serum. Although we could not find enterovirus RNA in breast milk, it is possible that the virus may be excreted into breast milk during peak viremia for a short period of time. In fact, such a phenomenon was recently described in 2 mothers whose infants experienced an enterovirus infection immediately after birth. These infections were confirmed to be caused by the same virus on the basis of 100% nucleotide sequence identity between the viruses isolated from the mother and child.

Enterovirus infections are common in children. This
was also confirmed in our study, where 43% of the subjects had their first infection by the age of 1 year and 73% by the age of 3 years. This is likely still an underestimation, because the large number of enterovirus serotypes makes it difficult to diagnose all infections, even with this kind of wide panel of tests. The current EIA methods are known to find ∼50% to 70% of infections.10,11 Our polymerase chain reaction (RT-PCR) assay detects practically all enterovirus types,14 but viral RNA is present in serum for only a short time (usually <2 weeks). As shown in our earlier studies, the proportion of infections diagnosed can be increased by combining antibody analyses (EIA) and the detection of viral RNA (RT-PCR).16 Enterovirus infections were also common in adults, because an enterovirus infection was diagnosed in 15% of the mothers during pregnancy. Enterovirus antibody levels decreased during pregnancy (Fig 2), which may be because of hemodilution (the plasma volume increases by 30%-50% during pregnancy22) or to changes in the maternal immune system.21 After pregnancy, the antibody levels regained their earlier levels within 2 to 3 months.

CONCLUSIONS

Our study suggests that breastfeeding confers protection against enterovirus infections, and that breast milk antibodies mediate a significant proportion of this protective effect. Accordingly, breastfeeding seems to be a feasible way to decrease the risk of enterovirus infections in young infants, particularly if they are immunocompromized or otherwise susceptible to severe infections.

ACKNOWLEDGMENTS

This research was supported by the European Commission DGXII, Contract BMH4-CT96-0233, the Sigrid Jusélius Foundation, the Juvenile Diabetes Foundation International (grants 192612 and 195003 to Dr Åkerblom and grants 197114 and 395019 to Dr Hyötty), the Academy of Finland, the Liv and Hälsa Foundation, the Novo Nordisk Foundation, the Pediatric Research Foundation in Finland, and the Finnish Diabetes Association.


REFERENCES


Adverse Associations of Infant and Child Sleep Problems and Parent Health: An Australian Population Study

Joanna Martin, MBBS, FRACP, Harriet Hiscock, MBBS, FRACP, MD, GradDip, Pollyanna Hardy, MSc, Belinda Davey, BSc, Melissa Wake, MD, FRACP, MBChB

ARTICLE

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

ABSTRACT

OBJECTIVE. Infant sleep problems are strongly associated with poorer maternal mental health. It is not known whether they are also associated with poorer paternal mental health, nor whether sleep problems in older children are associated with maternal or paternal mental health. We aimed to examine relationships between child sleep problems and maternal and paternal mental health and general well-being in each of the infant and preschool-aged groups.

METHODS. Participants of this cross-sectional survey included families of infants (n = 5107) and preschool-aged children (n = 4983) participating in the first wave of the nationally representative Longitudinal Study of Australian Children, surveyed March through November 2004. The primary outcomes were mother and father serious psychological distress (measured by the Kessler-6) and general health (parent report of general health taken from the 12-item Short Form Health Survey and dichotomized into poor versus good health). A primary caregiver’s report of the child’s sleep problem was dichotomized into moderate/severe versus none/mild.

RESULTS. The prevalence of severe psychological distress ranged from 3% to 5%, and prevalence of poor general health ranged from 8% to 11%. Moderate to severe sleep problems affected 17% of infants and 14% of preschool-aged children. Infant sleep problems were associated with poor general health in mothers and with poor general health in fathers. Preschool sleep problems were associated with poor maternal general health. In mothers with no past history of depression, infant sleep problems had a greater effect on severe psychological distress compared with mothers with a past history of depression.

CONCLUSIONS. Sleep problems are common in infants and preschool-aged children. Infant sleep problems, in particular, are associated with poorer health in both parents, especially the mental health of mothers with no past history of depression.
MENTAL HEALTH PROBLEMS are common in parents of young children. Postnatal depression (PND) affects ~15% of Australian mothers,1 similar to other developed countries.2 In fathers, the prevalence of depression in the postpartum period may reach 50% if the mother is also depressed.3 Prevalence in parents of preschool-aged children has received less attention, but in a community sample of 174 English parents of preschool-aged children, 18% of mothers and 6% of fathers scored in the clinical depression range on a validated PND screening measure.2 Depression impacts adversely on a parent’s mood, cognition, and ability to conduct day-to-day tasks.4 Maternal and paternal PND are associated with later child behavior problems5,6 and maternal PND with later child developmental problems.7

Maternal depression may result, in part, from sleep problems in children.8 In Australia, 46% of mothers of infants aged 6 to 12 months reported a problem with their child’s sleep8 as did almost 30% of parents of children aged 1 to 3 years, and 14% to 27% of parents of preschool-aged children.10 Mothers who report a sleep problem in their child are twice as likely to report clinically significant symptoms of depression,8,9 and improving problem sleep in infants also improves maternal mood.11 In preschool-aged children, sleep problems are associated with maternal report of greater limitations on daily functioning.8 However, the impact of sleep problems on fathers’ mental health and general health is unknown.

In addition to infant sleep problems, other risk factors for maternal PND include a past history of depression, depression in a partner,12 reduced partner and social support,13 marital conflict,12 and recent negative life events.14 Paternal PND risk factors include a past history of depression, employment in manual or working-class occupations,3,15 partner depression, and relationship dissatisfaction.16,17 Less is known about the risk factors for depression in parents of older children. In a study of depression in mothers of 2-year-old children, risk factors included depression at 8 months, less partner and social support, more negative life events, poor health, and having a “difficult” toddler.13 No study has examined risk factors for depression in fathers of preschool children.

Drawing on a large, representative, Australian sample, we therefore aimed to determine the relationship between (1) infant sleep problems and (a) maternal and paternal mental health and (b) maternal and paternal general well-being, and (2) preschool sleep problems and (a) maternal and paternal mental health and (b) maternal and paternal general well-being. We hypothesized that sleep problems would be common and would be associated with poor mental and general health in mothers and fathers, even after adjusting for other known maternal and paternal depression risk factors.

METHODS

Data were drawn from the first wave (March to November 2004) of the Longitudinal Study of Australian Children (LSAC), a national study targeting 5000 infants and 5000 4- to 5-year-old preschool-aged children. LSAC used a 2-stage cluster sampling design with Australian postcodes as the primary sampling units (stratified by state of residence and urban versus rural status) and children enrolled on the Medicare Australia database as the secondary sampling units. Families for whom a post office box but no street address existed were excluded, as were very remote postcodes, and only 1 child per family was included in the study. Of those who were resident in the sampled postcodes and contactable, response rates were 64% for infants (n = 5107) and 59% for preschool children (4983).

Procedures and Inclusion Criteria

As well as a face-to-face interview in the child’s home (conducted by trained researchers), both the primary and secondary caregivers completed written questionnaires that were distributed to them at the interview. Data for this article were drawn from both the face-to-face and written questionnaires. We included available data from mothers (“female primary caregivers”: 99% and 97% of infant and child primary caregivers, respectively) and fathers (“male secondary caregivers”: 89% and 84%, respectively). The study was approved by the Australian Institute of Family Studies Ethics Committee.

Exclusion Criteria

The small proportion of fathers who were primary caregivers was excluded from the study because their completed questionnaire did not include data regarding a past history of depression, which was considered a potential confounding factor.

Measures

The primary outcomes were maternal and paternal mental health and general well-being. Mental health was measured by using the Kessler-6 (K6), a validated, 6-item screen for psychological distress, with a score ≥18 indicative of serious psychological distress.18 General health was measured with the widely used single global health item from the 36-item Short Form Health Survey and the 12-item Short Form Health Survey (SF-12),19 in which parents rated the quality of their general health on a 5-point Likert scale (1 = poor, 5 = excellent). Responses were dichotomized into poor health (fair/poor) and good health (good/very good/excellent).20

The primary exposures were infant or preschool-aged child sleep problem. The primary caregiver (mothers) reported whether their child’s sleep was a problem (no, mild, moderate or severe problem). Responses were dichotomized into no/mild and moderate/severe problem.
Potential confounding variables of the relationship between child sleep problems and maternal outcomes included (1) a past history of depression (depressive symptoms lasting >2 weeks in the past year or depression experienced for at least the last 2 years), (2) partner support, determined by using a study-designed, single item asking “How often is your partner a resource or support to you in raising your child(ren)?” (1 = never, 4 = always); (3) overall support, determined by using a study-designed, single item asking “Overall, how do you feel about the amount of support or help you get from family or friends living elsewhere?” (1 = I get enough help, 3 = I don’t get any help at all; “I don’t need any help” was combined with the response “I get enough help”); (4) 2 or more stressful life events; and (5) overall relationship satisfaction (1 = extremely unhappy, 7 = perfect) determined by using a single item from Spanier’s Dyadic Adjustment Scale, a validated measure of relationship satisfaction.21 For children in the preschool-aged sample only, child temperament was included and was measured by taking the average score on each of the 3 subscales (Sociability, Persistence, and Reactivity) of the 12-item Short Toddler Temperament Scale22; an overall score of >1 SD above the mean was considered a “difficult” temperament and all others as an “easy/normal” temperament. For fathers, potential confounding variables differed in that serious psychological distress in their partner was included, but items regarding a past history of depression, overall support, and stressful life events were not included because these were not measured.

Potential sociodemographic confounding variables included number of caregivers in the home (mothers only), country of birth (Australia/New Zealand versus other, mothers of infants only), and employment status (fathers only).

Analysis
Separate analyses were conducted for children in the infant and preschool-aged cohorts. Children with and without missing data on the K6 and general health item were compared by using $\chi^2$ analyses on child gender, birth order, prevalence of sleep problems, and parental age and education level. To demonstrate differences in the key parent outcomes, we conducted $\chi^2$ analyses of the relationships between moderate/severe infant or preschool sleep problems and maternal/paternal (1) serious psychological distress and (2) poor general health. Unadjusted and adjusted odds ratios were calculated by using logistic regression for the effect of a moderate/severe infant or preschool sleep problem on maternal/paternal (1) serious psychological distress and (2) poor general health.

Tests of interaction assessed whether the impact of (1) an infant sleep problem on maternal serious psychological distress and general health was greater among mothers with a past history of depression than those without and (2) a preschool sleep problem on maternal serious psychological distress and poor general health was greater among mothers who reported a “difficult” child temperament than those who did not.

Analyses were weighted for the multistage sampling design, allowing for unequal probabilities of selection into the sample, and for nonresponse. First-order Taylor linearization was used to obtain estimates of standard error taking account of the correlation of responses within postcodes. Analyses were conducted by using Stata 9.1 (Statacorp, College Station, TX).

RESULTS
Sample Characteristics
A total of 75% to 85% of mothers and fathers who completed the K6 and general health item (see Fig 1) did not differ from the remainder in terms of parental age, child gender, birth order, or severity of sleep problem ($P > .05$). However, parents who completed the K6 and general health item were more likely to have completed secondary school than those who did not ($P < .001$). Table 1 outlines sample characteristics. Moderate/severe sleep problems were reported in 17% and 14% of infants and preschool-aged children, respectively. The prevalence of serious psychological distress ranged from 3% (fathers of preschool children) to 5% (mothers of preschool children), and of poor general health from 8% (mothers of infants) to 11% (fathers of preschool children). Compared with 2001 Australian census data, LSAC families were more likely to have a mother who had completed secondary education in infant (66.9% in LSAC vs 56.6% in census) and child (58.6% in LSAC vs 48.3% in census) cohorts. Children from lower income families were underrepresented (31.7% of infants in LSAC vs 41.2% in census, and 29.2% of preschool-aged children in LSAC vs 40.6% census, had a combined parental income of <$800 per week).21

Bivariate Analyses
In the $\chi^2$ analyses in mothers, there was strong evidence that both infant and preschool sleep problems were associated with serious psychological distress and poor general health ($P < .01$; Table 2). In fathers, infant sleep problems were associated with serious psychological distress and poor general health, whereas preschool sleep problems were associated with poor general health only ($P < .05$; Table 2).

Unadjusted and Adjusted Analyses: Infant Sleep Problems
Table 3 shows unadjusted and adjusted odds ratios for the effect of infant sleep problems on parent mental and general health. In mothers, on average, an infant sleep problem almost doubled the odds of serious psychological distress (unadjusted odds ratio [OR]: 1.76; 95% con-
fidence interval [CI]: 1.18–2.63) and poor general health (unadjusted OR 1.83; 95% CI: 1.39–2.41). After adjusting for potential confounding variables, this relationship persisted only for general health (adjusted OR: 1.50; 95% CI: 1.07–2.09). A past history of depression was the most important predictor of both psychological distress and poor general health (unadjusted OR: 19.08; 95% CI: 11.08–32.86). In the adjusted analysis, infant sleep problems seemed to have a greater effect on psychological distress in those without (adjusted OR: 4.58; 95% CI: 1.35–15.61) than those with a past history of depression (adjusted OR: 1.04; 95% CI: 0.62–1.75; P = .02 for test of interaction).

In fathers, an infant sleep problem increased the odds of poor general health (adjusted OR: 1.47; 95% CI: 1.11–1.94) whereas serious psychological distress in a partner increased the odds of serious psychological distress in the father (adjusted OR: 2.34; 95% CI: 1.13–4.85). Relationship happiness and partner support showed strong evidence of a protective effect on both the unadjusted and adjusted analyses, reducing the odds of serious psychological distress on average by almost half (P < .01).

**Unadjusted and Adjusted Analyses: Preschool Sleep Problems**

In preschool-aged children, relationships between sleep problems and parental mental and general health were weaker than in children the infant cohort (see Table 4). In the unadjusted analyses, sleep problems were 1 of 3 negative risk factors showing strong evidence of an association with maternal serious psychological distress and poor general health. The other 2 risk factors were past history of depression and stressful life events (P < .001). After adjusting, there was weak evidence of a relationship between sleep problems and mental and general health, (P = .06 and P = .05, respectively) whereas the other 2 negative risk factors continued to show strong evidence of a relationship (P < .001). There was no evidence that sleep problems in preschool-aged children had an impact on paternal mental health (adjusted P = .12) and borderline evidence of an association with general health (P = .06).

**DISCUSSION**

Infant sleep problems were common and were associated with serious psychological distress and poor general health in mothers and with poor general health but not serious psychological distress in fathers. Sleep problems in preschool-aged children were also common, and an adjusted analysis provides weak evidence of an association with poor maternal mental health and poor maternal and paternal general health.

This is the first study, to our knowledge, to examine in a nationally representative sample the associations between infant and child sleep problems and the mental health of fathers and the general health of both mothers.
The large sample size enabled calculation of precise estimates of parent-reported sleep problems and mental and general health, and well-validated measures were used for the primary outcomes.

The study has some limitations. First, only 59% of the eligible children in the preschool sample and 64% of children in the eligible infant sample took part in LSAC. Although weighting was used for all analyses to account for differences in nonresponders, this may limit generalizability to those population groups underrepresented in LSAC. Second, parents who did not complete the K6 or the general health question were less likely to have completed secondary school education and were more likely to report a lower household income; thus, the results of this study may not generalize to this group. Noncompleters may also have been more depressed but, if this were the case, the relationship between child sleep problems and parent mental health would have most likely strengthened had they participated. Third, because the data are cross-sectional, causal directions cannot be assumed. Fourth, the sleep measure was based on subjective parent report. However, parent report is an established marker of problematic child sleep patterns, and there are strong indications of the reliability of parent reporting. Finally, maternal report of child sleep could be influenced by depression. However, in a previous

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Characteristics of the Infant and Preschool Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>Infant Sample*</td>
</tr>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td>Parental Serious psychological distress</td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>4252</td>
</tr>
<tr>
<td>Father</td>
<td>3434</td>
</tr>
<tr>
<td>Poor general health</td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>4245</td>
</tr>
<tr>
<td>Father</td>
<td>3582</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td></td>
</tr>
<tr>
<td>Maternal</td>
<td>5032</td>
</tr>
<tr>
<td>Paternal</td>
<td>4552</td>
</tr>
<tr>
<td>Born in Australia/New Zealand</td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>5033</td>
</tr>
<tr>
<td>Father</td>
<td>4552</td>
</tr>
<tr>
<td>2-parent household</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5033</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>5029</td>
</tr>
<tr>
<td>Father</td>
<td>4475</td>
</tr>
<tr>
<td>Father’s employment</td>
<td></td>
</tr>
<tr>
<td>Skilled labor/not in labor force</td>
<td>3979</td>
</tr>
<tr>
<td>Unemployed/menial labor</td>
<td>556</td>
</tr>
<tr>
<td>Child Sleep problem</td>
<td></td>
</tr>
<tr>
<td>None/mild</td>
<td>5030</td>
</tr>
<tr>
<td>Moderate/large</td>
<td>17.1</td>
</tr>
<tr>
<td>Male</td>
<td>5033</td>
</tr>
<tr>
<td>English spoken at home</td>
<td>5030</td>
</tr>
<tr>
<td>Age, mean (SD), mo</td>
<td>5033</td>
</tr>
<tr>
<td>Birth weight, mean (SD), g</td>
<td>4999</td>
</tr>
<tr>
<td>Term delivery (&gt;36 wk)</td>
<td>5024</td>
</tr>
<tr>
<td>Assisted deliveryb</td>
<td>5029</td>
</tr>
<tr>
<td>First-born child</td>
<td>5033</td>
</tr>
<tr>
<td>Breastfeeding duration, mean (SD), wk</td>
<td>3035</td>
</tr>
</tbody>
</table>

* A total of 5033 mothers and 4553 fathers in the infant sample.
** A total of 4839 mothers and 4174 fathers in the preschool sample.
\(n\) is total number of mothers or fathers responding to each item in the questionnaire.
Data not collected for the preschool sample.
Cesarean section, vacuum extraction, forceps, or other.

and fathers. The large sample size enabled calculation of precise estimates of parent-reported sleep problems and mental and general health, and well-validated measures were used for the primary outcomes.

The study has some limitations. First, only 59% of the eligible children in the preschool sample and 64% of children in the eligible infant sample took part in LSAC. Although weighting was used for all analyses to account for differences in nonresponders, this may limit generalizability to those population groups underrepresented in LSAC. Second, parents who did not complete the K6 or the general health question were less likely to have completed secondary school education and were more likely to report a lower household income; thus, the results of this study may not generalize to this group. Noncompleters may also have been more depressed but, if this were the case, the relationship between child sleep problems and parent mental health would have most likely strengthened had they participated. Third, because the data are cross-sectional, causal directions cannot be assumed. Fourth, the sleep measure was based on subjective parent report. However, parent report is an established marker of problematic child sleep patterns, and there are strong indications of the reliability of parent reporting. Finally, maternal report of child sleep could be influenced by depression. However, in a previous
Community study of >600 families, we found that mothers who reported a sleep problem in their child also reported more frequent and longer night wakings and longer sleep onset delay, irrespective of maternal depression status. The prevalence of infant (17%) and preschool (14%) sleep problems is less than shown in previous Australian community samples (30% and 46%, respectively). This may result from a difference in how sleep problems were recorded, because our study classified those with a “mild” sleep problem as not having a problem. When we reclassified “mild” sleep problems as a problem in the children in the preschool cohort, the prevalence of sleep problems rose to a comparable 34%.

Similar to previous studies, we found that report of an infant sleep problem increased the odds of a mother reporting poor mental health. For mothers without a past history of depression, infant sleep problems played a major role in their mental health. However, mothers with a past history of depression also had high levels of psychological distress, but their child’s current sleep problem added little to this burden. For these mothers, genetic or biological factors may play a greater role in their mental health than external factors such as a child’s sleep. General health of mothers reporting infant sleep problems was also poorer. This is not surprising given the physical demands that motherhood entails which, when coupled with disturbed sleep or sleep deprivation, may lead to or exacerbate existing health concerns.

Sleep problems in preschool-aged children affected parents to a lesser degree. Mothers of preschool-aged children may adjust to their child’s sleep disturbance over time, or the difficulties may be related to settling to sleep (rather than waking overnight), thus disturbing parent sleep less. Sleep problems had little effect on fathers after adjusting for confounding variables, possibly because, as the authors’ clinical experience suggests, fathers play less of a role than mothers in managing the sleep problems of older children.

Given the high prevalence and co-occurrence of parent psychological distress and child sleep problems (particularly in infants), health professionals in regular contact with families should ask about both. This is important because of the known adverse impacts of depression on child health, development, and later behavior problems. If sleep problems are identified, effective approaches to their management should be offered. This may lead to improvement in parent mental health, especially in mothers with no past history of depression. Fathers should be actively engaged in the assessment and management of child sleep problems, because their health is also at risk. Future research should determine whether child sleep management strategies could also improve general health in mothers and fathers.

### ACKNOWLEDGMENTS

This article uses a confidentialized unit record file from the Longitudinal Study of Australian Children Project, which was initiated and is funded by the Commonwealth Department of Families, Community Services, and Indigenous Affairs and is managed by the Australian Institute of Family Studies. Dr Hiscock was supported by a Murdoch Childrens Research Institute part-time research salary grant and Dr Wake by National Health and Medical Research Council Population Health Career Development Award 284556 for the duration of this manuscript’s preparation.
### TABLE 3 Predictors of Serious Psychological Distress and Poor General Health in Mothers and Fathers of Australian Infants

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mothers of infants,</th>
<th>Fathers of infants,</th>
<th>Fathers of infants,</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Unadjusted OR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate/severe sleep problem</td>
<td>4249</td>
<td>1.76 (1.18–2.63)</td>
<td>.006</td>
</tr>
<tr>
<td>Past history of depression</td>
<td>4245</td>
<td>19.08 (11.08–32.86)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Partner support in raising children</td>
<td>3858</td>
<td>0.55 (0.45–0.67)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Overall social support/help</td>
<td>4275</td>
<td>0.45 (0.37–0.55)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>≥2 stressful life events</td>
<td>4261</td>
<td>3.50 (2.50–4.90)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Relationship happiness</td>
<td>3915</td>
<td>0.59 (0.53–0.67)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>2-parent household</td>
<td>4252</td>
<td>0.34 (0.22–0.53)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Born in Australia or New Zealand</td>
<td>4252</td>
<td>0.68 (0.47–0.98)</td>
<td>.04</td>
</tr>
<tr>
<td>Interaction</td>
<td>—</td>
<td>—</td>
<td>4.58 (1.35–15.61)</td>
</tr>
<tr>
<td>Sleep problem and no history of depression</td>
<td>—</td>
<td>—</td>
<td>1.04 (0.62–1.75)</td>
</tr>
</tbody>
</table>

---

a ORs adjusted for all other variables.

b The OR of parent having serious psychological distress compared with parent not having serious psychological distress. Departure from linearity for these ordinal variables was assessed, and in all cases the test result was nonsignificant at the 5% level. The reported ORs are linear effects representing multiplicative change in OR for each unit increase in the covariate (eg, for relationship happiness, the unadjusted OR of a mother having a serious psychological distress decreases by a factor of 0.59 [95% CI: 0.53–0.67] as relationship happiness increases by 1 point on a 7-point scale).

c Two-parent household was dropped from the regression because of too few observations in 2 of the 4 cells.

d P value for interaction.

e The interaction was not statistically significant at the 5% level for mothers’ general health, thus the interaction was dropped from the regression.


### Predictors of Serious Psychological Distress and Poor General Health in Mothers and Fathers of Australian Preschoolers

<table>
<thead>
<tr>
<th>Variable</th>
<th>Serious Psychological Distress</th>
<th>Poor General Health</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Unadjusted OR (95% CI)</td>
</tr>
<tr>
<td>Mothers of preschool-aged children, n</td>
<td>3456</td>
<td></td>
</tr>
<tr>
<td>Moderate/severe sleep problem(^b)</td>
<td>4085</td>
<td>2.23 (1.61–3.10)</td>
</tr>
<tr>
<td>Difficult child temperament</td>
<td>4071</td>
<td>1.48 (1.00–2.19)</td>
</tr>
<tr>
<td>Past history of depression</td>
<td>4075</td>
<td>18.79 (11.84–29.81)</td>
</tr>
<tr>
<td>Partner support in raising children(^c)</td>
<td>3536</td>
<td>0.63 (0.52–0.76)</td>
</tr>
<tr>
<td>Overall social support/help(^d)</td>
<td>4042</td>
<td>0.44 (0.37–0.54)</td>
</tr>
<tr>
<td>≥2 stressful life events</td>
<td>4058</td>
<td>4.78 (3.38–6.76)</td>
</tr>
<tr>
<td>Relationship happiness(^c)</td>
<td>3524</td>
<td>0.57 (0.51–0.64)</td>
</tr>
<tr>
<td>2-parent household</td>
<td>4087</td>
<td>0.28 (0.20–0.38)</td>
</tr>
<tr>
<td>Fathers of preschool-aged children, n</td>
<td>3100</td>
<td></td>
</tr>
<tr>
<td>Moderate/severe sleep problem</td>
<td>3173</td>
<td>1.55 (0.93–2.57)</td>
</tr>
<tr>
<td>Unemployed/ menial labor</td>
<td>3162</td>
<td>1.37 (0.71–2.63)</td>
</tr>
<tr>
<td>Serious psychological distress in partner</td>
<td>3137</td>
<td>3.28 (1.60–6.74)</td>
</tr>
<tr>
<td>Relationship happiness(^c)</td>
<td>3148</td>
<td>0.53 (0.42–0.66)</td>
</tr>
<tr>
<td>Partner support in raising children(^c)</td>
<td>3163</td>
<td>0.54 (0.42–0.71)</td>
</tr>
</tbody>
</table>

---

**a** ORs adjusted for all other variables in table.  
**b** The interaction was not significant for sleep problem and child temperament for either serious psychological distress or general health (P > .19), thus the interaction was dropped from both regressions.  
**c** OR of mother having a serious psychological distress compared with not having a serious psychological distress. Departure from linearity for these ordinal variables was assessed, and in all cases the test result was nonsignificant at the 5% level. The reported ORs are linear effects representing multiplicative change in OR for each unit increase in the covariate (eg, for relationship happiness, as relationship happiness increases by 1 point on a 7-point scale).  
**d** The top 3 categories of relationship happiness were collapsed into 1 category (extremely/fairly/a little unhappy) because of very small numbers of fathers with serious psychological distress in these categories.

---

**REFERENCES**

REPORT RAISES QUESTIONS ABOUT HIGH-SCHOOL COURSES

“American educators have complained about grade inflation for years. But new findings suggest that US high schools may also suffer from another type of inflation—in the labeling of courses. Under pressure to produce graduates better prepared for college and the workplace, dozens of states have stiffened high school graduation requirements in recent years, pushing a broader array of students to take more years of core subjects and eliminating less rigorous lower-tier courses altogether. Reflecting these efforts, a review of high-school transcripts by the staff of the National Assessment of Educational Progress shows that high school students are taking, and receiving higher grades in, more college-prep courses than ever. Yet just-released test results for 12th graders on the NAEP, a widely respected barometer of educational achievement known as the ‘nation’s report card,’ indicated that students are graduating with mediocre math skills and reading abilities that have tumbled to their lowest level since the early 1990s. The 12th-grade tests are designed to measure the sorts of high-level thinking demanded in college work. The findings raise questions about whether college-prep courses are as tough as their titles indicate, and, if so, whether high schools and their instructors are adequately prepared to teach such courses to a rapidly changing mix of students.”

Noted by JFL, MD
ABSTRACT

OBJECTIVE. Our goal was to estimate the effects of managed care program type on service use and access for publicly insured children with chronic health conditions.

METHODS. Data on Medicaid and State Children’s Health Insurance Program managed care programs were linked by county and year to pooled data from the 1997–2002 National Health Interview Survey. We used multivariate techniques to examine the effects of managed care program type, relative to fee-for-service, on a broad array of service use and access outcomes.

RESULTS. Relative to fee-for-service, managed care program assignment was associated with selected reductions in service use but not with deterioration in reported access. Capitated managed care plans with mental health or specialty carve-outs were associated with a 7.4-percentage-point reduction in the probability of a specialist visit, a 6.3-percentage-point reduction in the probability of a mental health specialty visit, and a 5.9-percentage-point decrease in the probability of regular prescription drug use. Reductions in use associated with primary care case management and integrated capitated programs (without carve-outs) were more limited, and integrated capitated plans were associated with a reduction in unmet medical care need. We failed to find significant effects of special managed care plans for children with chronic health conditions.

CONCLUSIONS. Managed care is associated with reduced service use, particularly when capitated programs carve out services. This finding is of key policy importance, as the proportion of children enrolled in plans with carve-out arrangements has been increasing over time. It is not possible to determine whether reductions in services represent better care management or skimping. However, despite the reductions in use, we did not observe a corresponding increase in perceived unmet need; thus, the net change may represent improved care management.
EXPERTS ARE MIXED in their expectations of how well managed care programs meet the needs of children with chronic health conditions (CWCHC). Although managed care holds promise for improved organization and accountability, questions have been raised about the capacity of managed care organizations (MCOs) to provide appropriate access to high-quality care. In particular, concerns have been raised about the breadth and adequacy of provider networks, and health plans’ lack of experience delivering care to CWCHC. Financial incentives to limit access to necessary but expensive services may be particularly problematic for children who have elevated service needs. Furthermore, enrollment of CWCHC in closed panel managed care systems could disrupt preexisting provider relationships.

The question of how Medicaid and State Children’s Health Insurance Program (SCHIP)–enrolled CWCHC fare under managed care is particularly important. CWCHC rely disproportionately on public insurance, and managed care is now the predominant financing and delivery mechanism for Medicaid and SCHIP enrollees, with an estimated 60% penetration in Medicaid by 2003. Most separate SCHIP programs have adopted managed care, further emphasizing its importance. Historically, children receiving Supplemental Security Income (SSI), a subgroup of CWCHC, were exempted from some mandatory managed care arrangements used in Medicaid programs, but the broader group of Medicaid- and SCHIP-enrolled CWCHC has been mainstreamed into managed care in most states. In recent years, SSI recipients in many states have also been enrolled on a mandatory basis in managed care plans.

Assessing the impact of managed care on CWCHC is complicated by the fact there is no single model of managed care. Rather, federal Medicaid policy has permitted a variety of program designs to emerge within Medicaid, and the advent of freestanding SCHIP plans resulted in the incorporation of an array of mainstream private MCOs into public programs. The experience of CWCHC may vary depending on how the delivery of care is organized and the nature of financial incentives to providers. Thus, to assess the extent to which the needs of CWCHC are being met, it is important to compare across different managed care program types.

In this article, we provide the first national estimates of the effects of different types of managed care programs on access to and use of health care services for publicly insured CWCHC. We examine not only the mainstream managed care programs, both those that integrate and those that exclude or “carve-out” mental health or specialty services, but also consider programs designed specifically for CWCHC. The results of our analyses provide important new information to state and federal policy makers concerned about access to high-quality care for this group of vulnerable children.

BACKGROUND AND PREVIOUS LITERATURE
Within the Medicaid and SCHIP programs, managed care programs generally fall into 1 of 2 major categories. Primary care case management (PCCM) programs designate a primary care physician who is paid a small monthly management fee and serves as the gatekeeper or manager of specialty referrals and other health care services. Capitated risk models pay a predetermined monthly fee to an MCO to arrange for provision of the full range of necessary services. The MCO may be a private health maintenance organization, or a risk-bearing provider group, often organized around traditional safety net hospitals and/or clinics. In a number of states, prescription drugs, dental care, mental health, substance abuse, and/or other specialty services are excluded from the principal MCO contracts, with services provided through a separate capitated network or on a fee-for-service (FFS) basis. In recent years, Medicaid and SCHIP programs in a small number of states and localities have designed managed care systems specifically for CWCHC. These systems are typically built around networks of primary and specialty providers that possess particular expertise serving CWCHC and are often paid risk-adjusted rates to account for the higher cost of serving this population. Enrollment in any of these programs may be mandatory or voluntary. Some programs may be mandatory for selected populations, whereas others may be included on a voluntary basis or excluded completely.

There is limited literature on managed care for publicly enrolled CWCHC. Previous research examined Medicaid programs in single states or counties by using pre-post designs or comparisons between SSI recipients, CWCHC not receiving SSI, and children without reported health problems. Results concerning access effects are mixed, with some studies finding that MCO enrollment reduces unmet needs, whereas other studies find the reverse. MCO enrollment in 2 studies was associated with reduced use of services and costs. However, the use of different study populations and relatively little information concerning the design of each program make it difficult to generalize from these studies or to compare the impact of different models of managed care.

In an effort to move beyond isolated state-specific studies and to explore the role of managed care program design on access and satisfaction, Hill et al used qualitative methods to compare the experiences of CWCHC served under FFS and 4 Medicaid managed care models in 8 states. Findings indicate that mainstream managed care systems provide high-quality primary and specialty medical care but often fall short in their ability to identify CWCHC and address their nonmedical needs. Mainstream capitated plans that carve out selected services caused considerable fragmentation in service delivery and were often confusing to families. Specialized managed care systems seemed to be the most successful in
meeting the diverse and complex needs of CWCHC. This study suggests mechanisms through which managed care plan design can affect access and use, and highlights the importance of carve-outs as a focus for additional analysis.

Variation in managed care program type across states and over time was used in several studies to compare effects of different types of Medicaid managed care plans, although not specific to CWCHC. For example, Garrett et al found that mandatory capitated programs reduced emergency department (ED) use and increased specialist visits for children. Mandatory programs where the family must choose either an MCO or PCCM increase the probability of having a usual source of care, reduce ED use, increase immunization completeness, and increase the probability of having a physician visit. Few effects were found for PCCM programs. However, the study did not examine outcomes separately for CWCHC; in fact, to achieve a more homogeneous population, children with SSI were dropped from the analytic sample. In our study, we extend the work of Garrett and colleagues but focus specifically on the effects of managed care plan types on access to care and use of services for CWCHC.

DATA AND METHODS
The primary source of data was the National Health Interview Survey (NHIS), an ongoing household survey that collects data on demographic characteristics, health status, insurance coverage, access to care, and use of health care services. Additional detail was captured for a sample child from each family, with an annual sample of 13,000. A knowledgeable parent or other adult reported on behalf of minor children. We pooled data from the person level and sample child files from 1997–2002, selecting children eligible for and enrolled in either Medicaid or SCHIP. We used the NHIS data to construct a series of child, parent, and family measures; identify CWCHC; and measure access and service use outcomes.

Identifying CWCHC on the NHIS
The definition of CWCHC is primarily diagnosis-based. We included children reported to have been diagnosed as having 1 of several common chronic conditions that appeared on a checklist on the NHIS. These conditions include attention-deficit/hyperactivity disorder, mental retardation, Down syndrome, asthma, cerebral palsy, sickle cell anemia, muscular dystrophy, autism, congenital or other heart disease, and diabetes. We also included children reported to have an activity limitation caused by a health condition lasting at least 12 months, to capture children with a condition not included on the checklist. We used information from a mental health scale to identify children reported to be unhappy or sad most of the time over the previous 6 months as a proxy measure for child depression. This definition is broader than the consequence-based definition of children with special health care needs adopted by the Maternal and Child Health Bureau. We chose not to use that definition because some of the factors used to identify affected children, for example, elevated service use, are outcomes for our study; thus, interpretation of the results might be confounded by changes in the population of interest. The final analytic sample included 13,550 children eligible for and enrolled in Medicaid or SCHIP, with one fifth of those identified as CWCHC.

Outcome Measures: Use of Services and Access to Care
We selected a variety of use and access measures from topics addressed in the NHIS. Indicators were created for any ED use; any inpatient stay; ≥10 health professional visits (designated as “elevated outpatient use”); any visit with a medical specialist, mental health specialist, or vision care provider, all during the past 12 months; and regular prescription drug use during the past 3 months. Indicators of access included the presence of a usual source of care other than a hospital ED and unmet needs for medical, dental, prescription drug, and mental health services.

Child, Parent, and Family Characteristics
In the multivariate models, we controlled for demographic characteristics of the child, parent, and family. Child characteristics included age, gender, race, ethnicity, and immigration status. We controlled for parents’ education, earnings relative to the federal poverty level, marital and health status, as well as family size, sibling health, and age. Because managed care program type may be correlated with local area characteristics, we also controlled for metropolitan statistical area size and county-level measures of primary and specialty physician supply as reported in the Area Resource File, the presence of an academic medical center as reported by the American Hospital Association, and private health maintenance organization penetration as reported by Interstudy. Fixed state effects were included to capture aspects of state policy that might affect outcomes but did not change over time, and year indicators were included to capture temporal trends that were independent of changes in managed care.

Data on Managed Care Plan Type
The NHIS collects information on managed care restrictions for persons enrolled in Medicaid, such as the need to get approval for referrals, but this information is inadequate to distinguish between different types of managed care. Instead of using the self-reported information on managed care enrollment, for each year of the NHIS we assigned information on the type of managed care implemented in the county of residence, specific to Medicaid and separate SCHIP programs, by CWCHC status. Although many states have implemented programs...
statewide, there are others where implementation is partial; thus, it is important to assign county-specific plan types. These measures and the process of linking managed care plan types to children are described in the following sections.

Information on managed care program type was abstracted from annual Medicaid managed care enrollment reports,23–25 special surveys of state Medicaid programs and SCHIP,17,26 and SCHIP state plans. The Enrollment Reports are submitted annually by states to Centers for Medicare and Medicaid Services, and describe the different types of managed care programs operating in the state, the covered services, the included and excluded populations, and for each plan, the counties in which it is operational. A county- and year-specific database was created on managed care type and whether enrollment was mandatory for children generally, and for SSI recipients enrolled in Medicaid or other CWCHC in SCHIP. Given the variety of plan types available, we grouped plans to make the analysis tractable. We split capitated programs based on the presence of mental health or specialty care carve-outs. The final managed care categories are FFS, voluntary plans, mandatory PCCM, mandatory “integrated” capitated plans (inclusive of mental health or specialty services), mandatory capitated plans with mental health or specialty service carve-outs, and specialized managed care systems for Medicaid-enrolled children with SSI or SCHIP-enrolled CWCHC. The latter group includes some programs that were optional for CWCHC, but the default program was a mandatory capitated plan.

Managed care data were linked to children eligible for Medicaid or SCHIP. Because many states had different managed care programs for Medicaid and separate SCHIP, we distinguished between children eligible for each program. We used a detailed algorithm that replicates the eligibility determination process, incorporating federal- and state-specific Medicaid and SCHIP eligibility rules. The algorithm models the application of most categorical, income, and resource tests, using data from the NHIS to create child or family level measures for each relevant eligibility comparison.

Specification of Access and Use Models
We estimated multivariate linear probability regressions for each dichotomous outcome measure. Because of the large number of control variables, we estimated models by using the full sample of eligible and enrolled children. The models included indicators for the types of managed care plans, with FFS as the reference category, an indicator for CWCHC, and interaction terms between them. The full effects for CWCHC, which we present in this article, are based on the sum of the main and interaction effects for which we test the significance by using Wald tests.26

The estimates should be interpreted as managed care program effects, reflecting the program reported by the state as being implemented in each county for each population group. This program effect may differ from the individual effect of being enrolled in a managed care plan, to the extent that not all the Medicaid or SCHIP enrollees are enrolled.17 With voluntary programs, there may be partial and disproportionate enrollment of selected groups, and even in counties with mandatory programs, some individuals may be exempted from the program and will opt out, whereas others may have difficulty enrolling in their managed care plan. In the results we present, program and individual enrollment effects should be reasonably close, because we limit our focus to the mandatory programs.

The public-use NHIS data do not include the state or county identifiers needed to link state Medicaid and SCHIP eligibility rules, or to link county-level measures of managed care type to individual observations. To access data files with these state and county indicators, we conducted all analyses at the National Center for Health Statistics’ Research Data Center in Hyattsville, Maryland.

All analyses were performed by using Stata 8 software. Sample proportions presented were weighted to national totals. Standard errors were adjusted to reflect the complex sample design of the NHIS. All results discussed are significant with \( P \leq .05 \), unless otherwise indicated.

RESULTS
CWCHC are enrolled in a variety of managed care programs, and the distribution has shifted over time away from FFS to various mandatory programs. The most prevalent types of managed care in the analytic sample are mandatory capitated programs with mental health or specialty carve-outs, affecting 25% of Medicaid- and SCHIP-enrolled CWCHC (Table 1). Integrated capitated programs are the next most common managed care type, affecting almost 22% of children. Only 6% of CWCHC were in counties with mandatory special programs, whereas almost one third (32%) were in counties with either FFS or voluntary managed care programs. The small representation of CWCHC in special programs is consistent with the small number of states and populations for whom these plans are relevant. In 1997, almost one quarter (24.5%) of CWCHC were enrolled in FFS programs. This proportion decreased to 11.1% by 2002. This shift was balanced by a dramatic increase in the proportion of CWCHC enrolled in capitated programs with carve-outs.

Consistent with their elevated needs, publicly insured CWCHC tend to have high levels of health care service use (Table 2). For example, almost one quarter had medical specialist (23%) and mental health specialist (24%) visits, 31% reported a vision care visit, and 38% reported regular use of prescription medications. Two in
TABLE 1  Distribution of Managed Care Plan Type Among Medicaid- or SCHIP-Eligible and -Enrolled CWCHC

<table>
<thead>
<tr>
<th>Managed Care Plan Type</th>
<th>All Years</th>
<th>1997</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFS</td>
<td>0.159</td>
<td>0.245</td>
<td>0.111</td>
</tr>
<tr>
<td>Voluntary PCCM, capitated or special program</td>
<td>0.157</td>
<td>0.168</td>
<td>0.114</td>
</tr>
<tr>
<td>PCCM program, mandatory</td>
<td>0.154</td>
<td>0.084</td>
<td>0.175</td>
</tr>
<tr>
<td>Integrated capitated, mandatory</td>
<td>0.218</td>
<td>0.247</td>
<td>0.227</td>
</tr>
<tr>
<td>Capitated with mental health or specialty carve-outs, mandatory</td>
<td>0.248</td>
<td>0.200</td>
<td>0.304</td>
</tr>
<tr>
<td>Special programs for CWCHC, mandatory</td>
<td>0.064</td>
<td>0.057</td>
<td>0.067</td>
</tr>
</tbody>
</table>

* Also includes programs with mandatory enrollment in either PCCM or capitated plan.

Source: Author’s analysis of 1997–2002 NHIS.

TABLE 2  Access to Care and Use of Services Overall and by Managed Care Plan Type for Medicaid- and SCHIP-Eligible and -Enrolled CWCHC

<table>
<thead>
<tr>
<th>Use and Access Outcomes</th>
<th>All</th>
<th>FFS Only</th>
<th>PCCM, Mandatory</th>
<th>Integrated Capitated Program, Mandatory</th>
<th>Capitated Program with Carve-outs, Mandatory</th>
<th>Special Program, CWCHC, Mandatory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service use, past 12 mo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elevated outpatient use</td>
<td>0.151</td>
<td>0.172</td>
<td>0.183</td>
<td>0.121b</td>
<td>0.129b</td>
<td>0.196</td>
</tr>
<tr>
<td>Any specialist visits</td>
<td>0.227</td>
<td>0.288</td>
<td>0.217a</td>
<td>0.201b</td>
<td>0.198b</td>
<td>0.240</td>
</tr>
<tr>
<td>Any mental health visit, age 2–17 y</td>
<td>0.242</td>
<td>0.246</td>
<td>0.191</td>
<td>0.261</td>
<td>0.218</td>
<td>0.315</td>
</tr>
<tr>
<td>Any vision care visit, age 2–17 y</td>
<td>0.307</td>
<td>0.341</td>
<td>0.295</td>
<td>0.329</td>
<td>0.293</td>
<td>0.271a</td>
</tr>
<tr>
<td>Dental visit w/in last year, &gt;2 y</td>
<td>0.750</td>
<td>0.735</td>
<td>0.689</td>
<td>0.745</td>
<td>0.773</td>
<td>0.793</td>
</tr>
<tr>
<td>Taken prescription for past 3 mo</td>
<td>0.384</td>
<td>0.409</td>
<td>0.394</td>
<td>0.399</td>
<td>0.338bc</td>
<td>0.454</td>
</tr>
<tr>
<td>Any ED visit</td>
<td>0.401</td>
<td>0.407</td>
<td>0.430</td>
<td>0.344a</td>
<td>0.399a</td>
<td>0.437</td>
</tr>
<tr>
<td>Any hospital stay</td>
<td>0.103</td>
<td>0.128</td>
<td>0.097</td>
<td>0.064bc</td>
<td>0.097</td>
<td>0.109</td>
</tr>
<tr>
<td>Access to care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has usual source of care, other than ED</td>
<td>0.951</td>
<td>0.954</td>
<td>0.953</td>
<td>0.937</td>
<td>0.961e</td>
<td>0.959</td>
</tr>
<tr>
<td>Unmet medical care need</td>
<td>0.035</td>
<td>0.046</td>
<td>0.036</td>
<td>0.026</td>
<td>0.037</td>
<td>0.041</td>
</tr>
<tr>
<td>Unmet prescription drug need</td>
<td>0.062</td>
<td>0.069</td>
<td>0.069</td>
<td>0.048</td>
<td>0.059</td>
<td>0.080</td>
</tr>
<tr>
<td>Unmet dental care need, age 2–17 y</td>
<td>0.095</td>
<td>0.116</td>
<td>0.107</td>
<td>0.085</td>
<td>0.088</td>
<td>0.082</td>
</tr>
<tr>
<td>Unmet mental health need, age 2–17 y</td>
<td>0.032</td>
<td>0.031</td>
<td>0.036</td>
<td>0.031</td>
<td>0.031</td>
<td>0.047</td>
</tr>
<tr>
<td>Unmet any need (medical, prescription, dental, mental health; age 2–17 y)</td>
<td>0.152</td>
<td>0.175</td>
<td>0.149</td>
<td>0.138</td>
<td>0.147</td>
<td>0.165</td>
</tr>
</tbody>
</table>

The difference between managed care type and FFS is significant at ^ 0.05 < P ≤ 0.10; a 0.01 < P ≤ 0.05; b 0.05 < P ≤ 0.10; and the difference between capitated plan with and without carve-out is significant at ^ 0.01 < P ≤ 0.05; c 0.05 < P ≤ 0.10.

Source: Author’s analysis of the NHIS, 1997–2002

5 reported at least 1 visit to a hospital ED, and more than half of those (22%) reported multiple visits (data not shown). One in 10 was hospitalized in the past year. Almost all (95%) had a usual source provider when sick. Among children 2 to 17 years old, 15% were reported to have foregone needed medical, dental, prescription drug, or mental health services; unmet need for dental care services was reported by almost 10%.

Comparison across managed care program types indicates that use levels for specific services, such as ≥10 health care provider visits and any specialist were lower in the 2 capitated programs compared with FFS enrollees. Children enrolled in integrated capitated programs had lower hospitalization rates (8.4% vs 12.8% in FFS), and CWCHC in capitated programs with carve-outs had lower rates of regular prescription drug use (33.6% vs 38.4% for FFS). These comparisons do not control for differences in characteristics across managed care plan types.

Estimated Effects of Mandatory Managed Care Plans

Figures 1 through 4 summarize the effects on use of services associated with the 4 groups of mandatory managed care plans, relative to FFS, after adjustment for child, family, and area characteristics, state, and year. Mandatory PCCM programs were associated with reductions in the probability of any specialist visit (−5.8 percentage points), mental health specialty visit (−5.0 percentage points), and hospital stay (−3.0 percentage points) and a reduction in reported unmet need for mental health services, but none of these estimates reached standard levels of significance.

Integrated capitated programs were associated with reductions in the probability of any specialist visits (−5.3 percentage points) and any ED visit (−5.8 percentage points), but only the estimated reduction in unmet need for medical care (−2.3 percentage points) reached the standard level of significance. When managed care plans carved out mental health or specialty services, there was
a reduced probability of specialist visits (−7.4 percentage points), mental health specialty visits (−6.3 percentage points), vision care visits (−8.2 percentage points), and regular prescription drug use (−5.9 percentage points). The negative effects on mental health visits and prescription drug use are significantly larger for plans with carve-outs relative to plans where these services are integrated. It should be noted that there is a high correlation between use of mental health and other types of service carve-outs, which may explain what seems to be a spillover effect on vision care and prescription drug services. Alternatively, differences in mental health specialty use may result in reduced use of prescription drugs.

In contrast to the other models, special managed care programs for CWCHC were not found to have significant effects on service use; small positive trends were observed but they do not meet standards of significance.

Special managed care programs for CWCHC were associated with an upward shift in physician visits, although it is not reflected in the measure of elevated outpatient use we report.

**Magnitude of Managed Care Plan Effects**

The findings reported reflect the marginal change in the probability of service use associated with each managed care model relative to FFS plans. Although some of these changes are relatively small, they represent substantial changes relative to the FFS base. For example, the 7.4-percentage-point reduction in specialist visits associated with capitated plans with carve-outs represents a 26% reduction relative to the 29% of children in FFS counties who visit a specialist.

**DISCUSSION**

Our results suggest that when CWCHC enrolled in Medicaid or SCHIP are assigned to managed care plans, they...
experience lower use rates of commonly needed services, relative to CWCHC enrolled in FFS. Consistent with the findings of Hill et al,\textsuperscript{16} the effects of managed care plans differed across type, with capitated programs with carve-outs presenting profiles that differed from integrated capitated programs, PCCM, and special programs for CWCHC. Although all PCCM and capitated programs reduced the likelihood of specialist visits to some extent, only capitated programs with carve-outs were associated with significant declines in the probability of any mental health provider use and reduced prescription drug use. The effects associated with PCCM enrollment, albeit weaker, were somewhat unexpected, because PCCM providers lack the strong financial incentives present in capitated plans to manage services. It is likely that PCCM providers also participate in capitated networks, and that practice patterns carry over from patients in one form of managed care to another. It is also possible that PCCM programs create barriers to specialty referral that may be similar to a capitated plan with carve-outs.

Mandatory special programs for CWCHC are designed with goal of facilitating access and improving care management. However, because we fail to find positive effects on perceived access or reductions in ED or hospital use, it is not possible to draw conclusions about how well these programs work. It may be that only the most severely affected children (ie, those with SSI) are assigned to these plans, or that small numbers in the study sample limit statistical power.

The difference in effects of capitated plans with and without carve-outs are fully consistent with other research addressing this issue.\textsuperscript{15} Service carve-outs are intended to steer children to more specialized provider networks, but if those networks are at financial risk for services, there is a disincentive to accept referrals. State-
specific studies of mental health carve-outs suggest that carve-outs encourage substitution of outpatient for inpatient care, medical for specialty mental health visits, and increased use of psychotropic medications.28–30 The findings from this study are consistent with the second mechanism, but the observed reduction in regular use of prescription drugs is inconsistent with the reported pattern of substituting medication use for specialty care.

The increasing reliance on capitated managed care programs with carve-outs may or may not be problematic, depending on whether the reductions in specialist, mental health, and vision care visits represent appropriate substitution of primary care providers and judicious use of specialists, or whether plans are providing inadequate service to CWCHC. If the latter situation were true, then we would expect to observe an increase in reported unmet need for services. Because such increases were not observed, these results do not support an interpretation that managed care plans were skimping on specialty services. More research is clearly needed to explore the incentives that plans establish for both primary care and specialty providers when carve-outs are present and how those incentives affect referrals for care between systems.

A few caveats related to this study should be noted. This analysis uses household survey data, which relies on parent report of child health status, use of services and access problems, along with child and family characteristics. Self-reported data are subject to reporting error, which may have affected selection of children into the sample of Medicaid- or SCHIP-eligible and -enrolled, assignment by chronic condition status, and reported outcomes. These errors in measurement, particularly for child health status, may result in downward bias of estimated effects. The study is also limited in that the NHIS access to care and service use measures are applicable to children generally. If managed care programs affect access to truly specialized services that are of particular value to CWCHC, this analysis would not capture those effects.

Although the study provides critical new information on managed-care effects for CWCHC, there is substantial heterogeneity within the group of children, and managed care may have differential effects on children depending on the nature of their condition and its severity and stability. Although it would be ideal to stratify children along those dimensions in this analysis, the NHIS health status measures are not likely to capture such nuances adequately.

It is important to note that the estimated effects represent national averages. Plans likely vary in how they are implemented across and even within states. This study opens the black box of managed care by exploring effects associated with carve-outs and special programs. Future research is ongoing to address whether explicit policy choices within capitated care plans are associated with different outcomes.

Much debate still surrounds the issue of whether managed care can work for CWCHC. This study does not resolve the debate, but does suggest that managed care is associated with changes in outpatient management for CWCHC without jeopardizing access. However, the large differential effects associated with carve-outs suggest a need for additional study, involving data with sufficient clinical detail to more clearly assess the adequacy and appropriateness of care delivered to CWCHC under managed care.

ACKNOWLEDGMENT
This research was funded by Maternal and Child Health Bureau grant R40 MC 02495-01.

REFERENCES

PHYSICIANS’ ATTIRE DOES NOT AFFECT PATIENT SATISFACTION

“Whether a doctor wears business attire, casual clothing or scrubs seems to make little difference in a patient’s satisfaction with treatment, the results of a new survey show. ‘This contradicts the long-standing belief that attire affects the level of patient comfort or the patient’s perception of physician competence and professionalism,’ lead author Dr Richard L. Fischer, from the University of Medicine and Dentistry of New Jersey in Camden, said in a statement. The findings, which appear in the American Journal of Obstetrics and Gynecology, are based on a satisfaction survey conducted over a 3-month period that included 1116 women who had an office visit with a new obstetrician-gynecologist. Unbeknownst to the patients, the 120 participating physicians were randomly assigned to wear business attire, casual clothing, or scrubs on a weekly basis. For more information, go to: http://health.yahoo.com/news/172225:_ylt-AhSi403H5Y7EruzH2FrUcbEqLeSf.”

Vermont Medical Society. News Scan. February 26, 2007
Noted by JFL, MD
"Urticaria Multiforme": A Case Series and Review of Acute Annular Urticarial Hypersensitivity Syndromes in Children

Kara N. Shah, MD, PhD, Paul J. Honig, MD, Albert C. Yan, MD

Section of Pediatric Dermatology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania

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

ABSTRACT

Acute annular urticaria is a common and benign cutaneous hypersensitivity reaction seen in children that manifests with characteristic annular, arcuate, and polycyclic urticarial lesions in association with acral edema. It is mistaken most often for erythema multiforme and, occasionally, for a serum-sickness–like reaction. Although these 3 entities may present in a similar manner, specific clinical features help to distinguish them, and it is important for the clinician to be able to differentiate among them. We present herein a series of 18 patients who were given a diagnosis of acute annular urticaria and review the clinical distinctions between acute annular urticaria, serum-sickness–like reactions, and erythema multiforme. Because of the frequency of its clinical confusion with erythema multiforme, we propose the term “urticaria multiforme” as a more apt description to highlight the distinctive clinical features of this urticaria variant.
Acute annular urticaria, an acute urticarial hypersensitivity syndrome, is a morphologic subtype of urticaria characterized by the acute onset of blanchable annular, arcuate, and polycyclic erythematous wheals (Fig 1). Associated angioedema of the face, hands, and feet is often encountered in affected patients (Fig 2). Dermatographism, the production of transient erythema and edema (“wheal and flare”) at sites of skin trauma, is common in this context and may manifest in linear or geometric patterns (Fig 3). Acute annular urticaria is reportedly more common in children 4 months to 4 years of age.1 Systemic symptoms are usually limited to fever of short duration (1–3 days) with or without other symptoms suggestive of a concomitant illness such as diarrhea or cough, and children are nontoxic in appearance. The eruption is self-limited, and episodes usually resolve within 8 to 10 days.

Although the clinical findings in a patient with acute annular urticaria are distinctive, the condition is often misdiagnosed as erythema multiforme or, less commonly, a serum-sickness–like reaction. Confusion arises when the urticarial lesions of annular urticaria display a dusky, ecchymotic center, which can be mistaken for the target lesion of erythema multiforme, or when the presence of fever and/or edema of the hands and feet misleads the clinician to diagnose a serum-sickness–like reaction. Although the clinical distinctions between annular urticaria and erythema multiforme have been highlighted previously, confusion still exists.2,3 To emphasize the distinctive clinical and morphologic manifestations of acute annular urticaria that can aid the clinician in differentiating acute annular urticaria from erythema multiforme, we propose the new term “urticaria multiforme.” We suspect that urticaria multiforme is underrecognized as a result of the paucity of reported cases in the literature and the lack of a clear, concise summary of the clinical features that distinguish urticaria multiforme from these other clinical entities. Here we describe our experience with 18 patients who were given a diagnosis of urticaria multiforme, many of whom were referred for consultation out of concerns for erythema multiforme, and delineate the important clinical features that distinguish between these 3 conditions.

Methods
A retrospective chart review of patients seen in consultation in both the inpatient and outpatient settings by the pediatric dermatology service at the Children’s Hospital of Philadelphia over a 4½-year period from August 2001 to April 2006 was approved by the Children’s Hospital of Philadelphia Institutional Review Board. Patients given a diagnosis of acute annular urticaria or urticaria multiforme were identified. Information obtained from the review of patient records included patient age and gender, antecedent symptoms including fever, documentation of recent immunizations, medication history, results of any diagnostic testing performed during the evaluation, and physical examination with a focus on the cutaneous examination and the presence or absence of angioedema and dermatographism.

Diagnostic criteria used in the diagnosis of urticaria multiforme are outlined in Table 1.

Results
Eighteen patients between 10 weeks and 17 years of age were given a final diagnosis of acute annular urticaria or urticaria multiforme (Table 2). Data on the prevalence of pertinent associated symptoms are presented in Table 3. The most common initial referring diagnosis was either “rash” or “erythema multiforme.” At the initial evaluation, most patients presented with 1 to 6 days of symptoms. A majority of the patients (12 of 18 [67%]) reported an antecedent upper respiratory infection, otitis media, or viral symptoms; fever was present in 8 patients (44%).

Although not performed for all patients, the results of complete blood counts were normal for all children except one who had an elevated white blood count and evidence of concomitant Mycoplasma infection (patient 1). The only other documented infections included 1
patient with streptococcal pharyngitis (patient 15) and 1 patient with adenovirus gastroenteritis (patient 3). Other laboratory evaluations, including Lyme serologies, Epstein-Barr virus serologies, erythrocyte sedimentation rate, and blood cultures, were obtained in selected patients, and their results were negative. Two patients (11%) had received routine childhood vaccinations before the onset of symptoms (Varivax in patient 11 and unspecified routine immunizations in patient 17). Synagis had been administered to 1 patient (patient 4) 36 hours before the onset of symptoms. Concurrent or recent antibiotic use was documented in 8 (44%) of 18 patients.

Typical features of urticaria and angioedema were observed in a majority of the patients. Pruritus was nearly universal and was reported in 17 (94%) of 18 patients. Hand and/or foot edema was seen in 11 (61%) of 18, and facial edema was seen in 11 (61%) of 18 patients; overall, either hand and/or foot edema or facial edema was reported in 13 patients (72%). Although not evaluated in all patients, dermatographism could be demonstrated in 8 patients (44%). None of the patients manifested true target lesions, skin necrosis or blistering, mucous membrane involvement, arthralgias, or arthritis.

The majority of patients with urticaria multiforme required combinations of systemic antihistamines, usually a combination of cetirizine, diphenhydramine, or hydroxyzine with or without ranitidine, to achieve satisfactory symptomatic relief. Three patients had been
started on systemic glucocorticoids by their primary care provider before their evaluation by the dermatology service. Systemic glucocorticoids were promptly discontinued in 2 patients. In 1 patient with a history of previous allergic hypersensitivity reactions that required systemic glucocorticoids, corticosteroids were tapered slowly over 1 week. Patient 6 had been admitted with a presumptive diagnosis of Kawasaki disease because of fever, rash, and peripheral edema and had received several doses of aspirin. We were consulted before initiation of intravenous immunoglobulin, because each dose of aspirin was accompanied by an exacerbation of the “polymorphous rash” in association with features of facial and peripheral angioedema. Aspirin is a known histamine-releasing agent and can exacerbate urticaria. Discontinuation of aspirin and administration of combination antihistamine therapy resulted in prompt resolution of the child’s clinical findings. Additional evaluation did not support a diagnosis of Kawasaki disease, and the child did not receive intravenous immunoglobulin. In all patients for whom follow-up was obtained, symptoms and signs remitted within 2 to 12 days.

## DISCUSSION

Urticaria multiforme, also known as acute annular urticaria or acute urticarial hypersensitivity syndrome, represents an allergic hypersensitivity reaction mediated predominantly by histamine and characterized by transient cutaneous erythema and dermal edema. It may be immunoglobulin E dependent or independent.

Urticaria multiforme is a distinctive morphologic form of urticaria that is often misdiagnosed as erythema multiforme or a serum-sickness–like reaction. Urticaria multiforme is a common presentation of urticaria in infants and children. Most patients in our series who were diagnosed with urticaria multiforme were infants or preschool-aged children between 2 months and 3 years of age (15 of 18 patients [83%]), with the youngest patient presenting at 10 weeks of age and the oldest at 17 years of age.

The diagnosis is typically made on clinical grounds and should not require skin biopsy. The individual lesions of urticaria multiforme, like typical lesions of urticaria, are evanescent, initially appearing as small urticarial macules, papules, or plaques, but they expand rapidly to form annular, arcuate, and polycyclic wheals that subsequently fade within hours. Centrally, lesions may display either central clearing or a dusky, ecchymotic, hemorrhagic hue, which has been reported to occur more commonly in infants with acute urticaria (up to 49% of infants aged 1–36 months). This dusky hemorrhagic hue resembles ecchymosis or purpura but

## TABLE 2

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Age</th>
<th>Gender</th>
<th>Antecedent Symptoms</th>
<th>Antecedent Infection</th>
<th>Antecedent Medication Use</th>
<th>Fever</th>
<th>Facial and/or Acral Edema</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17 y</td>
<td>Male</td>
<td>OM, bronchitis</td>
<td>Mycoplasma</td>
<td>Amoxicillin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>13 mo</td>
<td>Female</td>
<td>Rhinorrhea</td>
<td>—</td>
<td>Nitrofurantoin (prophylaxis)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2 y</td>
<td>Male</td>
<td>—</td>
<td>Adenovirus (stool)</td>
<td>—</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>11 wk</td>
<td>Male</td>
<td>—</td>
<td>—</td>
<td>Palivizumab (Synagis)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>9 mo</td>
<td>Female</td>
<td>OM</td>
<td>—</td>
<td>Augmentin</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3 y</td>
<td>Male</td>
<td>Vital</td>
<td>—</td>
<td>—</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>15 mo</td>
<td>Male</td>
<td>OM, URI</td>
<td>—</td>
<td>Mesalamine, 6-mercaptopurine, omeprazole</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2 y</td>
<td>Male</td>
<td>OM</td>
<td>—</td>
<td>Amoxicillin</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>15 mo</td>
<td>Male</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>7 y</td>
<td>Female</td>
<td>URI</td>
<td>—</td>
<td>Topiramate, cefprozil</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>12 mo</td>
<td>Female</td>
<td>—</td>
<td>—</td>
<td>Amoxicillin, immunizations (Varivax)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>10 wk</td>
<td>Male</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>2 y</td>
<td>Female</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>13 mo</td>
<td>Male</td>
<td>URI, diarrhea</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>8 y</td>
<td>Male</td>
<td>Pharyngitis</td>
<td>Streptococcal pharyngitis</td>
<td>Amoxicillin</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>17 mo</td>
<td>Male</td>
<td>URI</td>
<td>—</td>
<td>Albuterol</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>8 mo</td>
<td>Female</td>
<td>URI, OM</td>
<td>—</td>
<td>Amoxicillin, immunizations (unknown)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>18 mo</td>
<td>Male</td>
<td>URI</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

OM indicates otitis media; URI, upper respiratory infection; —, none.

## TABLE 3

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Prevalence, n/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pruritus</strong></td>
<td>17/18 (94)</td>
</tr>
<tr>
<td><strong>Angioedema</strong></td>
<td></td>
</tr>
<tr>
<td>Hands, feet</td>
<td>11/18 (61)</td>
</tr>
<tr>
<td>Face</td>
<td>11/18 (61)</td>
</tr>
<tr>
<td><strong>Angioedema of hands and feet or face</strong></td>
<td>13/18 (72)</td>
</tr>
<tr>
<td><strong>Dermatographism</strong></td>
<td>8/18 (44)</td>
</tr>
<tr>
<td><strong>Fever</strong></td>
<td>8/18 (44)</td>
</tr>
<tr>
<td><strong>Symptoms suggestive of recent viral or bacterial illness</strong></td>
<td>12/18 (67)</td>
</tr>
<tr>
<td><strong>Recent antibiotic use</strong></td>
<td>8/18 (44)</td>
</tr>
<tr>
<td><strong>Recent immunizations</strong></td>
<td>2/18 (11)</td>
</tr>
</tbody>
</table>
rapidly resolves with antihistamine or systemic corticosteroid therapy. Associated angioedema of the face, hands, and feet represents subcutaneous vascular leak with resultant dermal edema and has been reported to occur in 37% to 60% of patients with acute urticaria. This angioedema is self-limiting and has not been associated with laryngoedema.4–6 In our series, the presence of facial and/or acral edema was common and was documented in more than two thirds (72%) of the patients. Pruritus is also commonly seen in urticaria, with a reported prevalence of 60% to 89%, although excoriations are uncommon. Pruritus was an almost universal finding associated with urticaria multiforme that was seen in 94% of the patients in this study. Fever was much less common and was seen in only 4% of the patients in this study.

Many children with urticaria have a history of an antecedent viral or bacterial infection or recent use of a systemic medication, often an antibiotic. However, exhausts diagnostic evaluations for an infectious etiology are generally not helpful, and focused testing based on specific symptoms is advised. In our study, a history of an antecedent viral infection or use of a systemic medication, often an antibiotic, was reported in 37% to 60% of patients with acute urticaria. This angioedema is self-limiting and has not been associated with laryngoedema. In our series, the presence of facial and/or acral edema was common (72%) in the patients with laryngoedema. The angioedema is self-limiting and has not been associated with laryngoedema. In our series, the presence of facial and/or acral edema was common and was documented in more than two thirds (72%) of the patients. Fever was much less common (4%) of the patients in this study.

**![Figure 4](image) Erythema multiforme. Classic acral bull’s-eye target lesions with central necrosis are shown.**

**Table 4** Distinguishing Features of Urticaria Multiforme, Erythema Multiforme, and Serum-Sickness–Like Reactions

<table>
<thead>
<tr>
<th>Feature</th>
<th>Urticaria Multiforme</th>
<th>Erythema Multiforme</th>
<th>Serum-sickness–Like Reactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appearance of individual lesions</td>
<td>Annular and polycyclic wheals with central clearing or ecchymotic centers</td>
<td>Classis target lesion with annular lesions with purpuric or dusky, violaceous center (may blister), middle ring of palor and edema, outer ring of erythema or blisters</td>
<td>Polycyclic urticarial wheals with central clearing; may appear purpuric</td>
</tr>
<tr>
<td>Location</td>
<td>Trunk, extremities, face</td>
<td>Involvement of palms, soles common</td>
<td>Trunk, extremities, face, lateral borders of hands and feet</td>
</tr>
<tr>
<td>Duration of individual lesions</td>
<td>&lt;24 h</td>
<td>Days to weeks</td>
<td>Days to weeks</td>
</tr>
<tr>
<td>Fixed lesions</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Total duration of rash</td>
<td>2–12 d</td>
<td>2–3 wk</td>
<td>1–6 wk</td>
</tr>
<tr>
<td>Mucous membrane involvement</td>
<td>Oral edema common, no erosions or blisters</td>
<td>May see oral erosions or blisters of lips, buccal mucosa, and tongue; rarely involves conjunctivae, nasal, or urogenital mucosa; usually involving only a single site</td>
<td>Oral edema common, no erosions or blisters</td>
</tr>
<tr>
<td>Facial or acral edema</td>
<td>Common</td>
<td>Rare</td>
<td>Common</td>
</tr>
<tr>
<td>Dermatographism</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Fever</td>
<td>Occasionally, low-grade</td>
<td>Occasionally, low-grade</td>
<td>Prominent, high-grade</td>
</tr>
<tr>
<td>Associated symptoms</td>
<td>Pruritus</td>
<td>Mild pruritis or burning</td>
<td>Myalgias, arthralgias, lymphadenopathy</td>
</tr>
<tr>
<td>Common triggers</td>
<td>Antibiotics, immunizations, viral illness</td>
<td>Herpes simplex virus, other viral illness</td>
<td>Antibiotics</td>
</tr>
<tr>
<td>Treatment</td>
<td>Discontinue any new or unnecessary antibiotics or medications; combinations of H1 and H2 antihistamines may be helpful; systemic steroids can be helpful in more recalcitrant cases</td>
<td>Supportive care; early institution of systemic steroids can sometimes be helpful</td>
<td>Discontinue any new antibiotics or medications; H1 and H2 antihistamines; supportive care; consider systemic corticosteroids</td>
</tr>
</tbody>
</table>
only rare reports of urticaria seen in association with nonsteroidal antiinflammatory drugs such as ibuprofen. Food allergy has not been reported in association with acute annular urticaria in children. Children with urticaria frequently show an incomplete response to oral antihistamines and may require combination therapy. Patients with urticaria multiforme seem to benefit from the administration of systemic antihistamines, usually both an H1 antihistamine (e.g., diphenhydramine) and an H2 antihistamine (e.g., ranitidine). Systemic corticosteroids are rarely required except in the most severe cases, and we tend to avoid systemic corticosteroid administration when underlying infection is suspected unless patients remain symptomatic despite combination antihistamine therapy.

Urticarial multiforme is commonly misdiagnosed as either erythema multiforme or a serum-sickness–like reaction. Important clues in the history and clinical examination that help to differentiate between these 3 entities are outlined in Table 4. An important distinction is the fleeting duration of the lesions of urticaria multiforme, which usually last minutes to hours as opposed to the fixed lesions of erythema multiforme and serum-sickness–like reactions, which typically last from days to weeks. The presence of dermatographism, a transient, induced wheal-and-flare reaction that may be elicited by rubbing or scratching of the skin and which represents a mast cell–mediated cutaneous dermal hypersensitivity reaction to pressure, is common in children with urticaria multiforme but is not usually observed in erythema multiforme or serum-sickness–like reactions. Infants and children with urticaria multiforme also commonly manifest angioedema of the face, hands, and feet, which is not a feature of either erythema multiforme or serum-sickness–like reactions.

Erythema multiforme represents a cutaneous cytotoxic hypersensitivity reaction. Classically, erythema multiforme manifests as so-called “target” lesions characterized by a central dusky zone of epidermal necrosis, which may blossom into frank blisters, surrounded by an inner ring of pale edema and an outer ring of erythema (the classic “bull’s-eye” lesion) (Fig 4). Although true target lesions are not seen in urticaria multiforme, on some occasions the lesions of urticaria multiforme may appear somewhat dusky or ecchymotic in the center but do not develop frank necrosis, central or peripheral blistering, or crusting. These ecchymotic changes are evanescent and resolve within 24 hours. Herpes simplex virus is the most common etiology associated with erythema multiforme, although other systemic infections such as Mycoplasma pneumoniae and medications such as antibiotics have also been implicated as triggers of erythema multiforme.

Herpes simplex virus has not been identified as a causative agent in urticaria multiforme. Like urticaria multiforme, erythema multiforme is self-limiting and generally requires only symptomatic treatment.

In children with fever, urticaria multiforme may resemble a serum-sickness–like reaction. Both conditions can manifest with polycyclic urticarial eruptions and angioedema. True serum sickness is a systemic type III hypersensitivity reaction mediated by immunocomplex...
deposition and complement activation within blood vessels. It classically occurs 1 to 3 weeks after administration of animal serum or foreign proteins, is dose and frequency dependent, and resolves spontaneously without permanent sequelae within days to weeks. The characteristic cutaneous findings are fixed, polycyclic urticarial lesions, angioedema, and a serpiginous purpuric eruption on the lateral borders of the hands and feet (Fig 5). Systemic manifestations include vasculitis, nephritis with hematuria and albuminuria, arthropaligias and/or arthritis, myalgias, and lymphadenopathy. True serum sickness is very rare in children, because administration of animal serum or medications containing protein components occurs infrequently.

Serum-sickness-like reactions are much more common and are characterized by fever, arthropaligias, lymphadenopathy, urticaria, and angioedema. Immunocomplex formation and systemic involvement such as nephritis and vasculitis do not occur. Serum-sickness-like reactions in children have been reported most commonly in association with medications such as cefaclor, but have also been linked to buproprion, griseofulvin, and are characterized by fever, arthralgias, lymphadenopathy, urticaria, and angioedema. Immunocomplex formation and systemic involvement such as nephritis and vasculitis do not occur. Serum-sickness-like reactions in children have been reported most commonly in association with medications such as cefaclor, but have also been linked to buproprion, griseofulvin, monocyte, amoxicillin, sulfamethoxazole-tri- 

Serum-sickness–like reactions are much more common and are characterized by fever, arthropaligias, lymphadenopathy, urticaria, and angioedema. Immunocomplex formation and systemic involvement such as nephritis and vasculitis do not occur. Serum-sickness–like reactions in children have been reported most commonly in association with medications such as cefaclor, but have also been linked to buproprion, griseofulvin, minocycline, amoxicillin, sulfamethoxazole-tri-methoprim, penicillin, fluoxacillin, cefprozil, and carbamazepine. There are also postlicensure reports of serum-sickness–like reactions to the heptavalent conjugate pneumococcal vaccine. The treatment of serum-sickness–like reactions includes discontinuation of the offending agent, administration of systemic antihistamines, and administration of a 2- to 3-week course of systemic steroids for more severe symptomatic cases.

Clinicians who care for children should be able to recognize urticaria multiforme and differentiate this condition from erythema multiforme and serum-sickness–like reactions. A directed history and physical examination can reliably distinguish these conditions, which will help avoid unnecessary diagnostic testing and allow for appropriate treatment. Early in the course of the disease, it may be difficult to differentiate urticaria multiforme from its clinical mimics. As the course of the disease progresses, the correct diagnosis typically becomes clear. The transient nature of the urticarial lesions, the presence of dermatographism and acral angioedema in patients with urticaria multiforme, and a favorable response to combination antihistamine therapy with an H1-antihistamine and an H2-antihistamine within 24 to 48 hours will often aid in the correct diagnosis. The use of systemic corticosteroids should be reserved for more severe symptomatic cases. For children in whom an urticarial eruption persists or is associated with other systemic findings such as arthropaligias, persistent fevers, or abnormalities on routine laboratory evaluation, other diseases should be considered in the differential diagnosis. Other important differential diagnostic considerations are listed in Table 5.

**ACKNOWLEDGMENTS**

The Fig 5 image is courtesy of Lisa Zaoutis, MD.

**REFERENCES**

Therapy for Head Lice Based on Life Cycle, Resistance, and Safety Considerations

Mark Lebwohl, MD, Lily Clark, MD, Jacob Levitt, MD

Department of Dermatology, Mount Sinai School of Medicine, New York, New York

Financial Disclosure: Dr Levitt is vice president and a major shareholder of Taro Pharmaceutical Industries, Ltd, which manufactures and markets Ovide lotion. Drs Lebwohl and Clark have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

The timing of head lice maturation most favorable to their survival in the presence of anti-lice agents is the maximum time as an ovum (12 days) and the shortest possible time of maturing from newly hatched nymph to egg-laying adult (8.5 days). Pediculicides that are not reliably ovicidal (pyrethroids and lindane) require 2 to 3 treatment cycles to eradicate lice. Ovicidal therapies (malathion) require 1 to 2 treatments. Treatment with an agent to which there is genetic resistance is unproductive. In the United States, lice have become increasingly resistant to pyrethroids and lindane but not to malathion. Treatment with malathion has favorable efficacy and safety profiles and enables the immediate, safe return to school. Nit combing can be performed adjunctively. No-nit policies should be rendered obsolete.
Traditional pharmacological therapies for the human head louse, *Pediculus humanus* var. *capitis*, have focused on 1 or 2 courses of various ovicidal and pediculicidal topical therapies. Head lice, within the past 20 years, have developed resistance to nearly all first-line pharmacotherapy in the United States. The American Academy of Pediatrics recommends permethrin 1% as first-line treatment for head lice, a medicine for which resistance in the United States is extensively documented.1–4

Head lice infestations are not merely a nuisance. Untreated infection can lead to poor sleep and exorciation, which can occasionally become superinfected with methicillin-resistant *Staphylococcus aureus* (MRSA) or *Streptococcus*3–7 Social stigma, embarrassment, low self-esteem, and disgust often plague patients. Finally, from a purely functional aspect, many schools prevent children with nits from attending school. One study estimated that in 1998, 12 to 24 million days of school were lost secondary to no-nit policies.8 Such policies result in absenteeism, lost work for parents, missed education for the child, and needless anxiety. In addition, head lice carry a large economic burden. It is estimated that pharmacotherapy alone for head lice infestations costs the US economy up to $240 million per year. Estimates for combined direct and indirect costs may be as high as $1 billion per year.9

**LIFE CYCLE OF HEAD LICE**

Until recently, the life cycle of the human head louse has been difficult to quantify because of environmental variance and the unavailability of in vitro rearing systems. From a practical perspective, one can view the life cycle of *Pediculus humanus* var. *capitis* as follows: (1) egg without an eyespot (the eyespot indicating a developed nervous system); (2) egg from lay to hatch; and (3) first nymphal (instar) stage to egg-laying adult. Table 1 summarizes the time spent in each stage of the life cycle.5,10 Most lice treatments are traditionally pediculicidal and inconsistently ovicidal. In the context of a nonovicidal therapy, without considering pediculicide resistance, a “worst-case” scenario for therapy (or best case scenario for lice survival) is a life cycle with the longest time spent as an egg (12 days) and the shortest time spent as a non–egg-laying adult (8.5 days).

### GENERAL THERAPEUTIC CONSIDERATIONS

Considerations in evaluating a lice therapy must include an understanding of a therapy’s mechanism of action and resistance, prevalence of resistance, and safety. Application instructions also bear significance in light of the head lice life cycle. In the face of highly prevalent resistance to a particular molecule, therapy would likely be successful in only a small proportion of patients with lice infestations. Repeat treatments of resistant lice with preparations to which they are resistant will not kill the lice. In addition to wasting money, the patient is unnecessarily exposed to any associated toxicity of the therapy.

Putting aside the issue of resistance, a perfectly ovicidal and pediculicidal agent that acts on the louse nervous system requires 2 treatments separated at least 7 days apart. On day 0, all lice and eggs with eyespots would be killed. Those eggs without eyespots would develop eyespots by day 7 and thus be susceptible at that time. A solely pediculicidal agent would require 3 applications separated by 7 days (Fig 1). On day 0, all lice would be killed, leaving only newly laid eggs and eggs just about to hatch. Therapy on day 7 would kill those eggs that hatched. These nymphs would not have had time to mature to egg-laying adult. What would remain are those eggs that are 7 days old but did not hatch. Therapy anytime between days 13 and 15 would kill the nymphs from those eggs, precluding development to egg-laying adult.

Using average, rather than extreme, values for egg hatch (8.5 days10) and maturation time (9.7 days5,11) lowers the demand for treatment by a solely pediculicidal agent to days 0 and 9 (Fig 2). Indeed, so long as the average time to hatch is less than the average time to mature to egg-laying adult, only 2 treatments are theoretically necessary. The challenge is to identify the time interval after which all eggs should be hatched but before which new eggs are laid. The assurance of successful therapy is now at the mercy of favorable population statistics. Provided there is no resistance, enough lice may be killed to allow for stochastic extinction. These considerations become somewhat less relevant in the presence of resistant lice, which are not expected to respond regardless of treatment schedule.

### THERAPEUTIC OPTIONS

Lindane (γ-benzene hexachloride) noncompetitively inhibits the γ-amino butyric acid (GABA) receptor, which typically binds GABA, an inhibitory neurotransmitter.11 The neuronal hyperstimulation that ensues causes paralysis of the louse and eventually death secondary to inability to feed. Lindane’s efficacy has waned over the

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Female Head Lice Life Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage of Life Cycle</td>
<td>Minimum Time in Each Stage, d</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Egg without eyespot</td>
<td>0</td>
</tr>
<tr>
<td>Egg from lay to hatch</td>
<td>7</td>
</tr>
<tr>
<td>First instar to egg-laying adult</td>
<td>8.5</td>
</tr>
</tbody>
</table>

*No data on head lice are available; thus, 4 days is an estimate.*5
years, and it is inconsistently ovicidal. Because of its neurotoxicity, lindane carries a black box warning and is specifically slated as a second-line treatment by the US Food and Drug Administration (FDA). In addition, there is widespread resistance that renders this drug nearly obsolete for the treatment of lice. Lindane resistance is mediated by genetic mutations in the GABA receptor, resulting in decreased sensitivity of the receptor to GABA antagonists.

Lindane is labeled for a single use and only a 4-minute application time. According to the package insert, longer application times and multiple applications are deemed too risky because of potential neurotoxicity. This risk is perhaps overstated, in light of the peculiar discrepancy in labeling between the shampoo and the lotion, which is indicated for total body (neck down) application for up to 12 hours. Such labeling may be arbitrary or may be because of concerns about increased absorption from the scalp. With respect to head lice life cycle, a single application would not reliably eradicate lice. Even in the circumstance of no resistance, one would require at least 2, if not 3, treatments to eliminate an infestation (as it is considered pediculicidal only). Two and 3 consecutive treatments are deemed unsafe by the FDA.

Permethrin and pyrethrins, in this article referred to collectively as "pyrethroids," are the principal over-the-counter (OTC) pediculicides available in the United States. (With respect to nomenclature, technically, pyrethrins are natural compounds originating from Chrysanthemum cinerariaefolium, pyrethrins are the insecticidal component of pyrethrums, and pyrethroids are synthetic, rather than naturally occurring, forms of pyrethrins.) They affect voltage-gated sodium channels, causing delayed repolarization of the neuron by impeding sodium channel closure. Like lindane, these insecticides paralyze the louse through hyperstimulation of the nervous system, preventing it from feeding. Pyrethrins are a chrysanthemum extract. In practice, these products are overwhelmingly safe. Rare cases of asthma exacerbations and even death have been reported in individuals with ragweed allergy after using pyrethrins.
based products.19 This is clearly the exception rather than the rule. Piperonyl butoxide works synergistically with pyrethrins by inhibiting microsomal enzymes in the louse, preventing pyrethrin catabolism, thereby extending their action.23 Permethrin is a broad-spectrum synthetic pyrethroid that works similarly to pyrethrin.

Pyrethroid resistance is widespread. In the United States, it has been rigorously demonstrated in Massachusetts,2,3,23 Idaho,2 Florida,3,4,22 and Texas,22,23 The principal mechanism of pyrethroid resistance involves mutation of the α-subunit gene of the neuronal voltage gated sodium channel, conferring decreased sensitivity of the channel to pyrethroids (termed knock-down resistance, or the kdr allele).3,24,25 This mechanism indirectly renders piperonyl butoxide ineffective because this compound only serves to prevent degradation of a now ineffective pyrethroid. Additional resistance involves elevated glutathione S-transferase and monoxygenases, which serve to metabolize pyrethroids.26 If treatment failure is secondary to genetic resistance, as opposed to inappropriate application, using permethrin 5% should be no more efficacious than permethrin 1%.27

Genetic resistance to pyrethroids is widespread in both the Unites States and abroad, making these therapies increasingly less useful in treating head lice.2–4,22,23,28,29 At one point in time, efficacy rates for pyrethroids were high: in 1985, DiNapoli et al30 showed 96% efficacy and 62% efficacy of permethrin 1% and pyrethrins 0.3% with piperonyl butoxide 3% at 2 weeks after a single application; in 1986, Taplin et al17 showed 97% efficacy at 2 weeks with a single application of permethrin 1%; in 1986, Brandenberg et al14 showed 99% efficacy with permethrin 1% at 2 weeks with a single treatment; in 1987, Carson et al31 showed 100% efficacy and 93.5% efficacy of permethrin 1% and pyrethrins 0.3% with piperonyl butoxide 3% at 2 weeks after a single application; and in 1994, Bainbridge et al32 showed a 79.5% cure rate with permethrin 1% at 2 weeks using treatment at days 0 and 7. In 2004 and 2006, Meinking et al33 showed only 55% and 45% efficacy, respectively, with permethrin 1% at 2 weeks using treatment at days 0 and 7.

The best explanation for increased failure rates of permethrin is an increasing prevalence of resistant lice. A secondary explanation invokes the life-cycle analysis. According to the instructions for use of OTC pyrethroids, they are to be applied at days 0 and 7 to 10. Using the worst-case head lice life-cycle scenario, 3 applications would be required for these nonovicidal agents. Using the average life-cycle scenario, a second application would have to occur at day 9. It would seem that these product instructions are not consistent with the life cycle of the resistant head louse; however, in the absence of resistance, the products seem to be ovicidal and quite effective, sometimes with 1 treatment.

The instructions of OTC pyrethroid products place heavy emphasis on the adjunctive performance of nit combing in achieving efficacy. Of course, with perfect nit combing, pharmacologic therapy is not needed. Unfortunately, numerous studies and observations prove nit combing success rates to be far from perfect.5,36,37

Malathion (derived from Latin and Greek, meaning “bad sulfur,” referring to this compound’s smell) is an organophosphate insecticide. In the louse, malathion is converted to malaoxon, which irreversibly inhibits acetylcholinesterase. The excess cholinergic activity causes neuronal hyperexcitability, thereby preventing the ability to feed. Resistance to malathion, when it occurs, is thought to result from increased levels of carboxylesterases that metabolize malathion into nonmalaoxon intermediates and from decreased sensitivity of acetylcholinesterase to malathion and malaoxon.38–40 Additional mechanisms are esterase sequestration of malathion and elevated metabolism by cytochrome P450 monoxygenases, glutathione S-transferases, and phosphotriesterases. Malathion resistance has not been reported in the United States. It has proven to be the most efficacious head lice therapy, outperforming the pyrethroids and lindane in vitro and permethrin in the clinic.4,12,34,35 Malathion was reintroduced into the US market at the request of the Centers for Disease Control and Prevention in 1998 because of concerns of pediculicide resistance.41

Only 1 pharmaceutical preparation containing malathion exists in the United States. Its high degree of efficacy has been attributed not only to the presence of malathion 0.5%, but to the presence of a high concentration of isopropyl alcohol (78%) and terpineol (12%) in its excipients. Isopropyl alcohol has demonstrated clinical activity in head lice.42 It is a nonspecific toxin that likely denatures louse proteins in the egg and adult. It also may serve to dehydrate eggs.4 Terpineol is a tea tree oil extract that has been demonstrated to have ovicidal and pediculicidal activity.43 Terpineol both inhibits acetylcholinesterase and binds octopamine receptors, causing neuronal hyperactivity and death in insects.44,45

Efficacy is attributed to the triple action of malathion with isopropyl alcohol and terpineol, likely making this a resistance breaking formulation. The probability of simultaneously evolving resistance to all 3 active agents is small. Similar approaches to infectious agents are found with multidrug antituberculosis regimens and highly active antiretroviral therapy for the human immunodeficiency virus. Of interest, British lice resistant to local malathion products were killed by the US formulation.46

Malathion, in its US formulation, is both ovicidal and pediculicidal.47 Approximately 80% of patients are cured with a single application, the remainder requiring a sec-
ond application 1 week later. Because malathion is both ovicidal and pediculicidal, its labeling of 1 to 2 treatments separated by 7 to 9 days properly correlates with the head lice life cycle.

Unfortunately, many misconceptions abound regarding the safety of malathion in an isopropyl alcohol vehicle. Because malathion lotion is safe, the FDA has intended it for first-line use in head lice. In the United Kingdom, malathion is available OTC for the treatment of head lice. The United Kingdom Committee on Safety of Medicine takes the position that “there is no evidence to suggest that serious systemic adverse reactions are associated with topical malathion.” Reports of accidental ingestion of malathion are exceedingly rare. Actually, between 1998 and 2003, malathion had fewer recorded symptomatic cases (<50) from unintentional ingestion than lindane (>700) or the pyrethroids (>300). The presence of a child-proof cap, a small orifice for egress of medication from the bottle, and an extremely unpleasant odor would seem to deter ingestion.

Concerns about flammability seem to be ill-founded, because there are no known reports of bodily injury resulting from the isopropyl alcohol catching fire. Appropriately, there are no reports of bodily injury from heat sources during use likely account for the lack of reports of flammability. Such precautions should not be misconstrued as a reason not to prescribe the product. With respect to malathion itself, a 59-mL bottle of malathion 0.5% lotion contains 295 mg of malathion. Transdermal absorption was minimal (between 0.2%–3.2%) when malathion 0.5% in isopropyl alcohol was applied to the scalp for 10 hours. This corresponds to a very small amount (9.44 mg at most) of malathion. Furthermore, no clinical effect was observed on plasma or erythrocyte cholinesterase activity levels when malathion in isopropyl alcohol was applied to the scalp. With respect to any malathion that is absorbed, it is rapidly metabolized by tissue A-esterases and carboxylesterases to inactive metabolites (mainly malathion monocarboxylic and dicarboxylic acid) that are subsequently excreted in the urine. Only with overwhelming doses of malathion does the liver metabolize malathion to its active metabolite, malaoxon. This is in contrast to other organophosphate insecticides, where metabolism in humans is not so efficient, resulting in poor selective toxicity. Of note, past reports of malathion toxicity rest largely with the use of agricultural grade malathion. In the 1970s, there were reports of “epidemic” malathion toxicity resulting in several deaths and several hundred ill. It was discovered that the material in question contained toxic impurities, including isomalathion, a potent noncompetitive inhibitor of carboxylesterase that prevented malathion detoxification. The malathion currently manufactured for topical application is extremely pure and of significantly lower concentration than that used in agriculture.

Resistance to first-line topical agents has encouraged some to search for alternative, off-label treatments for head lice. Two such therapies are oral ivermectin and oral trimethoprim/sulfamethoxazole. Ivermectin causes an influx of chloride ions across neuronal membranes resulting in paralysis in many types of parasites. This therapy is only pediculicidal because for lice to be exposed, they must take a blood meal that contains the drug. Lice feed every 4 to 6 hours. Minimum blood levels of ivermectin required to kill a louse have not been delineated. Without this knowledge, rational dosing with respect to the head lice life cycle cannot be made. Dosing for head lice follows accepted dosing for scabies, which is 200 μg/kg. Head lice life cycle dictates 3 treatments to ensure lice eradication. Were blood levels to become subtherapeutic in 1 day (where 1 half-life is 18 hours), therapy at day 0 would eradicate all lice and any eggs hatched on days 0 and 1. Repeat treatment would be required on day 10 to prevent a day-2 hatchling from maturing to egg-laying. This leaves eggs that were laid just at the time of initial therapy to hatch at day 12. A third treatment anytime between days 13 and 20 would guarantee clearance of lice in the absence of resistance. Resistance to ivermectin has not been reported to date in head lice, and its mechanism is poorly understood in other arthropods. Ivermectin is contraindicated in children who weigh <15 kg as it can cross the blood brain barrier. Its use has been suggested to be efficacious for head lice infestations resistant to topical therapies; however, predictably, 1-time dosing was not effective.

Trimethoprim/sulfamethoxazole is presumed to work by ridding lice of symbiotic bacteria in their gut. The lice presumably die from the lack of B vitamins that the bacteria synthesize; however, Meinking asserts that whereas body lice symbiotes are capable of B-vitamin synthesis, head lice symbiotes are not, calling into question this entire mechanism of action. One small study concluded that monotherapy with either trimethoprim or sulfamethoxazole is ineffective. Sulfamethoxazole carries the risk of Stevens Johnson syndrome, and in 1 trial for head lice, 3 (4%) of 76 patients experienced allergic rash. Dosing is 10 mg/kg per d based on trimethoprim, given in divided doses, because the drug’s half-life is 10.1 hours. Trimethoprim/sulfamethoxazole is not ovicidal. The proposal of a rational dosing regimen based on the head lice life cycle is complicated by the lack of data on how long an individual louse would survive and be able to lay eggs during continuous exposure to trimethoprim/sulfamethoxazole blood meals. The efficacy of trimethoprim/sulfamethoxazole in treating head lice is controversial at best, with 1 large study showing no benefit and another showing significant benefit.
Nonpharmacologic approaches involve occlusion therapy, nit combing, and hair removal. None has been proven, and many have been disproven, to be effective enough to be considered viable therapies. Vinegar, mayonnaise, petroleum jelly, olive oil, butter, isopropyl alcohol (70% by volume), and water submersion up to 6 hours fail as effective “occlusion therapy” to eliminate an infestation, being neither completely pediculicidal nor ovicidal.63 Lice do not have air sacs or lungs but, rather, obtain air by both diffusion and air in channels that tunnel throughout their bodies. When threatened with suffocation, lice can use spiracles to occlude air tunnels without suffocating and can survive for prolonged periods without air. Successful suffocation can only be achieved by blocking 100% of the louse’s spiracles, as well as the entirety of its cuticle. In addition, lice have a pressure mechanism that allows them to open blocked spiracles.64

Hair removal (that is, shaving the head) has never been formally or rigorously proven but is anecdotally effective.65 Because the lice require hair shafts to lay eggs, removing the hair should theoretically prevent the lice from propagating. Although effective, the cosmetic result is less than desirable, especially for school-aged girls.65,66 Combing with a fine-toothed comb has relatively low cure rates as monotherapy. One large trial demonstrated that nit combing yielded a cure rate of 38%,36 whereas another large study reported a cure rate of 57%.67 The former faired worse than malathion applied at days 0 and 7,36 whereas the latter fared better compared with malathion or permethrin single-use therapy. The studies were performed in the United Kingdom, in areas where resistance to both permethrin and malathion exist.36 Viable nits are well camouflaged and often so close to the scalp that a nit comb cannot effectively reach them. Furthermore, those who claim efficacy with nit combing note it must be performed rigorously for many minutes over many days, which is not practical for most people.36,37,68 In any event, it is clear that nit combing has some merit but perhaps adjunctively rather than as monotherapy.

In addition to administering direct pharmacological and/or nonpharmacological therapy for head lice, some environmental treatments might be considered as a secondary measure. Head lice that have fallen off the scalp are believed to be too weak to reinfest; however, no study has either proven or disproven this presumption.57 Lice survival away from the host is 6 to 26 hours, at which point they die from dessication and/or starvation.5 In light of the above, routine house cleaning, including vacuuming of floors and furniture and laundering of linens and clothing, is recommended. Spraying of furniture and mattresses with permethrin spray is not recommended. Heat (hot wash and hang clothes dryer) killed head lice experimentally placed in pillowcases, whereas cold wash and hang pillowcases out to dry did not kill head lice. Therefore, for those fomites (that is, hats, combs, brushes, stuffed animals, bed linens, clothing, etc) that are launderable, washing in water >50°C to 60°C or placing for 15 minutes in a hot clothes dryer is recommended.59,70 Grooming tools and toiletries should not be shared. Because nymphs will die if they do not feed immediately after hatching, it is unlikely that head lice will be propagated by a fomite harboring an egg. Put in another way, during the 12-day window during which an egg might hatch, it is more likely to be away from the host than on the host, such that a blood meal would be unavailable at the time of hatching. In light of the above argument, the extremist might store those items that are nonlaunderable in a closed plastic bag for 13 days, but this is not recommended.

The primary mode of head lice transmission is direct head-to-head contact.5,37,71 The prevention of new infestations and reinfections must be considered as part of head lice management. In addition to and likely more important than the environmental measures taken above, is the identification of head lice in close contacts. Close contacts have been as narrowly defined as bedmates and as broadly defined as household members and classmates. Detection of head lice and nits, especially in light infestations and by those unfamiliar with head lice, is notoriously imperfect. Many infestations are missed by visual inspection, and many cases of head lice are asymptomatic.5 One method to improve detection is via the nit combing of wet hair and examining the threads of the comb.72

The American Academy of Pediatrics, in 2002, published the following guidance in this journal: “If an index case is identified, all household members should be checked for head lice . . . .” The implication is that close contact, which in our opinion, include classmates, should be screened. On the basis of the findings of Williams et al,73 18 (31%) of 91 children with nits had concomitant lice, and 19 (18%) of 50 of those followed with nits alone converted to active infestation. On a theoretical level, if head lice is not a self-limiting infection, and if a child has not been previously treated for lice, then having “nits alone” would seem to imply that either the person screening is not trained in the detection of live lice or the screen was a false-positive. On the flip side, “progression to lice” must result either from hatching nits or reinfection from a close contact.

The argument for not screening is that screening classmates would result in the exclusion of children from school. The issue here is not one of whether to screen or not to screen (and screening should take place), but what school policy should be with respect to a positive screen. Indeed, there is no way to identify which of the 18% of children with nits alone will develop an active infestation or which are false-positives for nits or false-
negatives for live lice. On a theoretical basis, given the negligible risk and high efficacy, one approach is to treat all such patients simultaneously with malathion. Indeed, a study is warranted to validate this strategy with this agent. The alternative is to accept a continual, baseline lice prevalence in schools.

Because identification is imperfect, because head lice are known to exist in the close contacts of actively infected scalps, and because of the excellent safety profile of some available therapies, strong arguments can be made for the systematic treatment of close contacts regardless of screening. Indeed, the simultaneous treatment of contacts is the most reliable way to eliminate head lice outbreaks in households, schools, and communities and has been repeatedly validated.\(^5,37,74\) Although such a strategy is good in theory, many parents, nurses, and physicians may be averse to exposing those individuals not definitively infested to treatment. A compromise might be to encourage the treatment of household contacts and require at least careful screening of classmates of an index case by the school nurse.

A rigorous cost analysis of treatment with 1 or another agent is beyond the scope of this article. However, some general statements can be made. Malathion costs about $125 per bottle, lindane costs about $125 per bottle, and OTC permethrin costs approximately $10 to $15 per bottle. Looking at cost of therapy alone, permethrin is only 3 to 5 times cheaper than malathion considering the need for 2 to 3 bottles of the former, in the absence of OTC resistance. Malathion is covered by most Medicaid and managed care plans, and the copay is roughly equivalent to the out-of-pocket expense of an OTC. OTC failure, which is common, carries with it the economic burdens of additional physician visits (that often add to out-of-pocket costs), repeat purchases, lice transmissibility (raising treatment costs exponentially), missed work by parents (resulting in lost wages), and loss of state funding from schools because of student exclusion from school. The collective cost of OTC failure relative to that of malathion success would seem to exceed greatly the differential in product price. Most importantly, the child would be free from infection with malathion yet still have the infection with permethrin.

The out-of-pocket expense that families are willing to bear to eliminate head lice infestation is illustrated by professional nit-picking services available in some communities. They can charge up to $250/hour for the first 2 hours and $90/hour for each additional hour of nit picking. In addition, they offer a variety of nit combs, “chemical-free” anti-lice hair tonics, and informational materials that cost collectively up to $95.\(^75\) A household with >1 infested child could easily expend $1000 to avail themselves of these services.

CONCLUSIONS

In light of the review of therapeutic modes of action, resistance considerations, and head lice biology, malathion, specifically in the formulation containing isopropyl alcohol and terpineol, is the favored first-line agent for head lice. A patient infected with head lice at any given time will have lice existing at different points in the life cycle. The only therapy that, when used according to the package insert, will ensure head lice eradication is malathion (Table 2). Regardless of adequate therapy, efforts are fruitless if the population in question is resistant, which is currently a concern for both lindane and pyrethroids and not for malathion as available in the United States.

---

**TABLE 2**  
Current FDA-Approved Pharmacologic Therapies for Head Lice

<table>
<thead>
<tr>
<th>Medication</th>
<th>Brand Name</th>
<th>Treatment Recommendation on Package Insert</th>
<th>Treatment Recommendation Based on Life-Cycle Considerations</th>
<th>Safety Risks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lindane 1%</td>
<td>Formerly, Kwell; now, marketed as Lindane shampoo</td>
<td>1 application</td>
<td>Day 0, day 7, and once within days 13 to 15(^a) Days 0 and 9(^b)</td>
<td>Neurotoxicity FDA black-box warning: not to be used in patients with psoriasis, atopic dermatitis, or those &lt;110 lb Second-line therapy No more than 1 application Pregnancy category: C Pregnancy category: B</td>
</tr>
<tr>
<td>Permethrin 1%</td>
<td>Nix</td>
<td>Day 0, then repeat in 7 d if evidence of active infestation</td>
<td>Day 0, day 7, and once within days 13–15(^a) Days 0 and 9(^b)</td>
<td>May cause asthma attack if allergic to ragweed Pregnancy category: B</td>
</tr>
<tr>
<td>Pyrethrins 0.33% plus Piperonyl butoxide 4%</td>
<td>Pronto Plus, Rid, A-200</td>
<td>Day 0, then 7 to 10 d later</td>
<td>Day 0, day 7, and once within days 13–15(^a) Days 0 and 9(^b)</td>
<td></td>
</tr>
<tr>
<td>Malathion 0.5%</td>
<td>Ovide</td>
<td>Day 0, then 7 to 9 d later if evidence of active infestation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Using the worst-case scenario for life cycle: 12-day incubation, 8.5-day maturation to egg-laying adult.  
\(^b\) Using the average values for life cycle: 8.5-day incubation, 9.7-day maturation to egg-laying adult.
The use of malathion as a first-line treatment also has broader implications for school head lice policy in the United States. Currently, no-nit policies can exclude children from school unnecessarily, as is the case when only nonviable nits are present on the scalp. Conversely, insofar as nits represent the possibility of infestation and detection of active infestation is imperfect, the need for a more definitive demonstration of freedom from lice exists. A possible answer is to require index cases of head lice and their family members to be treated with malathion. Classmates could be screened for head lice and those found to have head lice treated promptly with malathion, preferably simultaneously and preferably at days 0 and 7. Rescreening after malathion treatment would not be necessary, because the examination is imperfect and the probability of treatment success is very high. As was done with Nix in the past, a child could return to school the following day with proof of treatment. Were the concern that a parent is neglectful or would not apply the product correctly, direct observational therapy could be used in the school. Such a practice would allow for the elimination of no-nit policies without risk of reinfection in the school and with all the attendant pharmacoeconomic benefits: breaking the cycle of spread, decreasing absenteeism from schools, decreasing missed work by the parent, and eliminating repeated spends on ineffective modalities.

REFERENCES

34. Meinking TL, Vicaria M, Eyerdam DH, Villar ME, Reyna S, Suarez G. Efficacy of a reduced application time of Ovide lotion (0.5% malathion) compared to Nix creme rinse (1% permethrin) for the treatment of head lice. Pediatr Dermatol. 2004;21:670–674
35. Meinking TL, Vicaria M, Eyerdam DH, Villar ME, Reyna S, Suarez G. A randomized, investigator-blinded, time-ranging study of the comparative efficacy of 0.5% malathion gel vs Ovide lotion (0.5% malathion) or Nix creme rinse (1% permethrin) used as labeled, for the treatment of head lice. Pediatr Dermatol. 2007; In press
47. Lwegaba A. Shaving can be safer head lice treatment than other pesticides. J Biochem Mol Toxicol. 2005;19:257–261
BLIND PEDESTRIANS SAY QUIET HYBRIDS POSE SAFETY THREAT

“For blind people, crossing the street is becoming even more of a challenge. Michael Osborn, a blind marketing consultant from Laguna Beach, Calif., and his guide dog, Hastings, were in the middle of an intersection one morning last April when the yellow Lab stopped short. Mr Osborn took the cue and halted—just in time to feel the breeze from a car passing right in front of them. ‘Half an inch and it would have hit us . . . . It wasn’t making any noise,’ says Mr Osborn, 50, who has been blind for 12 years. Witnesses say the car was a Toyota Prius, a hybrid vehicle. Hybrids deliver better mileage and less pollution than traditional cars by switching between a gasoline engine and an electric motor. But when operating on the electric battery, especially when idling at a stop or running at low speeds, the engine in a hybrid is almost silent. A hybrid vehicle is generally quieter than a vacuum cleaner. ‘I’m an environmentalist, and I’m all for quiet cars,’ says Mr Osborn. ‘But it poses a particular problem for somebody who has no vision.’ Blind pedestrians using a guide dog or cane are largely dependent on the sounds of traffic to cross streets safely. For a blind person, ‘it’s very important to be able to gather auditory and tactile cues from the environment,’ says Sumara Shakeel, of Toms River, NJ, who is a rehabilitation teacher for the New Jersey Commission for the Blind.”


Noted by JFL, MD
Pain Reduction During Pediatric Immunizations: Evidence-Based Review and Recommendations

Neil L. Schechter, MD\textsuperscript{a,b,c}; William T. Zempsky, MD\textsuperscript{b,c}; Lindsey L. Cohen, PhD\textsuperscript{d}; Patrick J. McGrath, PhD\textsuperscript{e,f}; C. Meghan McMurtry, BA\textsuperscript{e,f}; Nancy S. Bright, BSN\textsuperscript{a,c}

\textsuperscript{a}Pain Relief Program, Connecticut Children’s Medical Center, Hartford, Connecticut; \textsuperscript{b}Department of Pediatrics, University of Connecticut School of Medicine, Farmington, Connecticut; \textsuperscript{c}Department of Pediatrics, St Francis Hospital and Medical Center, Hartford, Connecticut; \textsuperscript{d}Department of Psychology, Georgia State University, Atlanta, Georgia; \textsuperscript{e}Centre for Research in Pediatric Pain, IWK Health Centre, Halifax, Nova Scotia, Canada; \textsuperscript{f}Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada

Financial Disclosure: Dr Zempsky is a consultant to Aresiva Pharmaceuticals.

\section*{ABSTRACT}

The pain associated with immunizations is a source of anxiety and distress for the children receiving the immunizations, their parents, and the providers who must administer them. Preparation of the child before the procedure seems to reduce anxiety and subsequent pain. The limited available data suggest that intramuscular administration of immunizations should occur in the vastus lateralis (anterolateral thigh) for children <18 months of age and in the deltoid (upper arm) for those >36 months of age. Controversy exists in site selection for 18- to 36-month-old children. A number of studies suggest that the ventrogluteal area is the most appropriate for all age groups. Longer needles are usually associated with less pain and less local reaction. During the injection, parental demeanor clearly affects the child’s pain behaviors. Excessive parental reassurance, criticism, or apology seems to increase distress, whereas humor and distraction tend to decrease distress. Distraction techniques vary with the age, temperament, and interests of the child, but their efficacy is well supported in the literature. Sucrose solution instilled directly into the mouth or administered on a pacifier reduces evidence of distress reliably in children <6 months of age and should be used routinely. Although there is no perfect topical anesthetic available at this time, selective use for children who are particularly fearful or who have had negative experiences in the past is highly endorsed. Pressure at the site, applied with either a device or a finger, clearly reduces pain. Finally, in the era of multiple injections, it seems that parents prefer that multiple injections be given simultaneously, rather than sequentially, if there are enough personnel available. Immunizations are stressful for many children; until new approaches are developed, systematic use of available techniques can significantly reduce the burden of distress associated with these procedures.
The development and administration of immunizations are among the greatest public health achievements of the 20th century, and their positive impact on disease prevention and reduction of human suffering is almost incalculable. Hundreds of millions of cases of illness and millions of deaths have been prevented by these agents. To provide this protection, the current Centers for Disease Control and Prevention schedule recommends immunizations against 14 diseases, which translates into 14 to 20 separate injections before the age of 2 years, depending on the number of combination vaccines available. Therefore, immunizations are the most frequently occurring painful procedures performed in pediatric settings. The number of immunizations now recommended necessitates that multiple injections be administered at the same visit (eg, 3 separate injections at the 2-month visit and 3 at the 4-month visit).

Despite the proven benefit of these procedures, the pain associated with these injections is a source of great anxiety and distress for many. In fact, recent research suggests that not only the children themselves but also their families and health care providers have concerns about the pain associated with multiple frequent injections.

For children, concerns about injections are often evident at the beginning of any clinical encounter. Although modern pediatric practice has broadened in scope to include health supervision and counseling, many children are so preoccupied with the possibility of an injection that this worry dominates the entire visit and limits the opportunity for other interventions. Every nurse or physician who works with children has entered an examining room to find a cowering child whose first question is, “Am I going to get a shot?” The needle is a powerful negative symbol for many children, and it is a phobia for some, and unfortunately has become iconic of medical care. All children, even at young ages, seem to have a pain memory, to anticipate painful procedures, and to react more intensely if they have had previous painful procedures with inadequate analgesia. There are enormous differences among children in their responses to injections, however, and it is clear that a host of variables, both inherent and environmental, may amplify or dampen their responses to injections.

Family members also are often quite concerned about immunization pain. Meyerhoff et al, in an attempt to quantify parental concern regarding multiple immunizations, developed a “willingness-to-pay” method for estimating that distress. According to their survey of 294 families drawn from a random sample of 26 centers around the United States, parents reported they would be willing to pay an average of $57 to avoid a 2-injection visit and nearly $80 to avoid a 3- or 4-injection visit. Regardless of the specific numerical sum parents reported and the veracity of their actual willingness to pay it, the data of Meyerhoff et al indicate significant parental concern about the pain associated with immunizations. That conclusion is supported by other research that suggests that parental concern about injections affects their compliance with medical care.

Similar unease exists among providers. Woodin et al reported that, depending on the child’s age, 56% to 65% of practicing physicians had strong concerns about 3 injections at a single visit. Eighty percent had strong concerns about administering 4 injections at a single visit, irrespective of the child’s age. In fact, practicing physicians were statistically more likely to have concerns about multiple injections at a single visit than were parents. Similarly, Reis reported that doctors and nurses were 6 times less likely to give all of the immunizations at a given visit if a child was scheduled for $3 injections than if he or she was scheduled for $2.

Despite the frequency with which immunizations are administered and the distress they provoke in children, their parents, and the providers who must administer the immunizations, there is a paucity of evidence-based data available on many aspects of good immunization practice. The literature in this area is often anecdotal, based on experience rather than research, and scattered through journals of the various disciplines that are involved in immunization administration. As a result, there is presently no available evidence-based algorithm that addresses all of the strategies that have evolved to reduce the pain of immunizations in children of all ages.

Because members of the Pain Relief Program at Connecticut Children’s Medical Center have had long-standing interest in this topic, we applied for and received a grant from the Mayday Fund to host a consensus conference to more formally review the existing literature and to generate from it recommendations regarding best practices for reduction of injection pain. Initially, we separated the immunization process into multiple components (eg, preparation, technique, location, needle length, and local anesthetic use). We then identified experts (defined as individuals who had published research in these areas) and invited those who were available to participate in the conference. This panel was supplemented by research methodologists and experienced local clinicians who were actively involved in immunizations in their practices. Relevant literature was distributed to all participants before the meeting. At the meeting, each expert presented a review of his or her assigned area, offered recommendations, and detailed the evidence base supporting those recommendations. When there were randomized, clinical trials or meta-analyses with clear results, the task was straightforward. When the evidence was limited, the group discussed the available information until a consensus was reached. This approach allowed for recommendations that were supported by the literature and bolstered by informed opinion. The results of those deliberations are summa-
ized here, and a clinical algorithm of best practices for immunization pain reduction is offered.

It is important to note that there are a number of factors that cannot be modified that affect the pain associated with immunizations. Age, gender, temperament, previous painful experiences, and cultural background all contribute to the child’s pain experience; obviously these factors cannot be altered, but they can be recognized as moderators of the experience of pain. Some of these factors have been examined in relation to immunization pain, whereas others have not. A comprehensive discussion of these variables is significantly beyond the scope of this review but is available elsewhere.

For purposes of discussion, the injection process has been divided into 2 time periods, that is, before the injection and during the injection. Aspects of the immunization before the injection that are reviewed include preparing the child and family, site selection for the injection, selection of needle length and gauge, and specific properties of the injectate. Elements during the injection itself that are reviewed include parental demeanor, use of sucrose, use of topical anesthetic agents, nonpharmacologic and physical strategies, and specific aspects of administration technique.

BEFORE THE INJECTION

Preparation for Immunization

Preparing the child and his or her family for a painful procedure has long been considered appropriate practice, and data support the argument that better preparation results in decreased distress and improved adjustment for the parent and patient. A number of studies have shown that adequate preprocedural preparation can reduce anxiety and procedural pain for a range of medical events, including venipuncture, dental procedures, surgery, and voiding cystourethrography. Although there is a rich literature examining preparation for various medical events and hospitalizations, there are few studies examining preparation for pediatric immunizations. Two experimental studies that explored preparation for immunization produced inconclusive results. Eland found no difference in pain ratings reported by children who were given information about the impact of a vapocoolant spray on immunization pain and those who were not. Cohen found that a coping skills training video viewed before immunization was ineffective in either increasing coping or decreasing child distress behavior (accounting for 1% of the variance in both) during the immunization.

In a survey study, Megel et al examined how parents prepared their children before preschool immunizations. Five types of preprocedural preparation/discussion were postulated: information sharing (what will happen), sensory information (how it will feel), justifying the procedure (explaining why the procedure is necessary), teaching relaxation strategies, and role playing. The results suggested that parents used a mixture of various types of preparation. Seventy-five percent of children received informational preparation from their parents, typically involving a description of the events that would occur. Of the 25% of children who received no information, 9 children were <3 years of age. Forty-two percent of parents also used some sensory information in their description. Forty percent of parents offered a rationale for receiving the injection. Relatively few parents offered any strategies for how to cope with the procedure (eg, relaxation, breathing, or distraction). Unfortunately, the relationship between the type of preparation and the child’s subsequent distress was not reported by the researchers.

Although there are few studies of preparation for pediatric immunizations, the rich literature on preparation for other painful procedures and hospitalizations allows some recommendations to be made. For parents, information about the immunization should include the reason for the vaccination, with emphasis on its benefits, and realistic information about how much pain is associated with the immunization procedure. In addition, parents should be provided with techniques (breathing techniques, reading or telling a story, or involving the child in a fantasy) they can use to “coach” their children through the procedure, to promote coping.

For children, the extensiveness and style of preparation should be guided by each child’s age and developmental level. In general, specific discussion about the immunization itself has more relevance for children >2 years of age. Research suggests that preparation should have at least 3 components, namely, what will happen (where, how long it will last, and what will be done), how it will feel (pressure, temperature, and level of discomfort to be expected), and strategies to cope with the stressor. Children should be asked what strategies they think will help them to cope and, if possible, those strategies should be incorporated into the immunization administration. In addition, given the strong data supporting distraction, that technique should be used during the immunization procedure.

In terms of the timing of preparation, literature findings are mixed. Studies have found that providing information too far in advance or too close to the event can heighten anxiety. Although some research revealed optimal timing for particular age ranges, those studies were not examining preparation for immunizations. Extrapolating from the literature on the development of children’s sense of time, it seems that toddlers and preschool-aged children should be informed of the injection as close to the actual administration as possible, to prevent escalating anxiety. For older children, much depends on their individual coping styles. In summary, findings from the broader preparation literature suggest...
that preparation for pediatric immunization is important and there are some indications regarding how it might be performed, but clear detailed guidelines are not yet available, because of the limitations of research on preparation specifically for pediatric immunizations.

Injection Site Selection
Despite limited supporting evidence, there seems to be agreement among most major professional and educational organizations in this field regarding the sites of intramuscular injections. At this time, most endorse the use of the anterolateral thigh (vastus lateralis) for infants and the upper arm (deltoid) for children >18 months of age.1,38–41

The anterolateral thigh was selected for young infants because of its relatively large muscle mass and lack of vital structures. The shift from thigh to arm should occur when the upper arm has adequate muscle mass to allow injection. This shift is driven by research that suggests that injection in the thigh is more painful and causes more incapacitation than injection in the arm. Ipp et al42 found that, at 18 months, severe pain occurred in 30.5% of patients injected in the thigh but only 8.1% of those injected in the arm. Those authors stated that 50% of patients injected in the thigh had decreased movement of the extremity, compared with 35% of those injected in the arm. Two thirds limped for 24 to 48 hours after the immunization. Ipp et al42 also noted, however, that redness and swelling occurred significantly more frequently in the deltoid group. They suggested, on the basis of their data, that the shift to the arm should occur at 18 months. Others, such as the County of Los Angeles Department of Public Health40 and the Colorado Immunization Project,43 suggested that 36 months is a more-appropriate changeover time. The Red Book of the American Academy of Pediatrics1 is nonspecific on this issue. It suggests that parents and toddlers prefer the arm at 18 months, whereas some health care professionals prefer the thigh through toddlerhood.

It should be noted, however, that there is an alternative view to the one presented above. For the past 20 years, the nursing literature has suggested that an alternative location, the ventrogluteal site, is preferred for both younger and older children.44–47 The ventrogluteal site, colloquially known as the “hip site,” is identified by “placing the palm of your hand over the greater trochanter, index finger over the anterosuperior iliac tubercle, and middle finger along the posterior iliac crest. The needle should be injected perpendicularly into the center of the V formed by the separated fingers.”46 Cook and Murtagh46,47 reported on 2 studies in which children 2 to 18 months of age receiving either whole-cell pertussis vaccine or acellular pertussis vaccine were assigned randomly to receive their injection in either the anterolateral thigh or the ventrogluteal site. The ventrogluteal site had lower rates of systemic reactions (irritability and persistent crying) and local reactions (bruising and erythema). Parental acceptability for the ventrogluteal site was greater in both studies. For reasons that are unclear, despite nursing enthusiasm, this site is not endorsed by any of the major medical organizations, although it is mentioned as an option in the recent edition of the Red Book.1

Use of the buttocks (gluteus maximus), which had been common in the past, is no longer supported, for a variety of reasons. Sciatic nerve injuries were reported in the past when the buttocks were used for intramuscular injections, although Thompson49 argued that those injuries did not occur with modern medications. In fact, MacDonald and Marcuse50 solicited, from the practicing community, case reports of children who had sciatic nerve injury secondary to vaccine injection in the buttocks, and only 1 case was reported to them. In addition, the buttocks lack adequate phagocytic cells, which are necessary for antigenicity; therefore, there may be a delay in the development of the immune response51 if they are used. Because of these concerns, the buttocks are no longer used for intramuscular injections except when large volumes of injectate are necessary, such as for immunoglobulin.52

Although the evidence is limited and somewhat controversial, it is generally agreed by influential stakeholders in the area, such as the American Academy of Pediatrics, that the anterior thigh (vastus lateralis) should be used for intramuscular immunization injections until at least 18 to 36 months. After that, the deltoid muscle is suggested. In rare situations in which a very large volume of injectate is necessary, the buttocks may be considered. The ventrogluteal site typically is not included in recommendations that originate from medical authorities. Clearly, however, additional research comparing the various sites with respect to pain, local reaction, antigenicity, and parental acceptability is necessary. Only then can recommendations that are supported by adequate evidence be generated.

Needle Length
The length of the needle selected for intramuscular injections has been examined in a number of studies. Although it would seem intuitive that the shortest needle with the thinnest gauge would produce the least trauma and pain, this does not seem to be the case. A number of studies support the contention that longer needles, which are more likely to penetrate muscle than are shorter ones, cause less pain and fewer adverse effects.

Diggle and Deeks53 examined 4-month-old infants scheduled to receive their diphtheria-pertussis-tetanus (DPT)/Haemophilus influenzae type b vaccine in the anterolateral thigh. The patients were assigned randomly to receive the immunization with either a 16-mm (5⁄8-inch), 25-gauge needle or a 25-mm (7⁄8-inch), 23-gauge
needle. The nurses used the World Health Organization injection technique, in which the skin is stretched taut and the needle is inserted at a 90° angle. Parents used a diary to report on the redness, swelling, and tenderness at the site for 3 days after the injection. More than one half of the infants vaccinated with the 16-mm needle developed redness and swelling initially, whereas only one third of the group vaccinated with the longer needle did. This difference between the groups persisted and in fact increased over the 3 days of observation. Differences in the reported rates of tenderness were not significant. An unfortunate flaw in the study was the fact that needle length was not the only independent variable, because the gauge of the needle also was changed. A meta-analytic review of this subject by Davenport, however, supported the conclusions drawn by Diggle and Deeks. The previously mentioned study by Ipp et al of 18-month-old children receiving DPT and polio vaccines also examined needle length. Those authors identified greater redness and swelling when immunizations were administered into the thigh with a smaller (16-mm) needle than with a longer (25-mm) one.

Two other studies call these conclusions into question, however. Studies by Cook and Murtagh and Groswasser et al both used ultrasonographic techniques to measure the subcutaneous tissue and muscle layer thickness of 2-, 4-, 6-, and 18-month-old children. Both studies identified the fact that, if the World Health Organization technique of immunization (skin taut, 90° angle) was used, then the shorter needle was adequate; however, if the skin was bunched, then the 16-mm needle deposited drug in the subcutaneous tissue and not the muscle. Diggle, responding to the article by Goldwasser et al, identified the fact that most children in that study were smaller than the normal population (most were between the 10th and 50th percentiles), which casts doubt on the generalizability of the finding that the smaller needle was adequate. In addition, the ultrasound studies were only indirect measures of potential pain and irritation and not direct measures. Complexity is added with the suggestion by Zuckerman that wider-gauge needles are associated with less localized erythema because they dissipate the injectate over a wider area.

The available data, therefore, are somewhat confusing. The Royal College of Pediatrics and Child Health, in response to the lack of clarity, thought that there was insufficient evidence to recommend any change in practice. It can be said, however, that, for larger infants or with skin-bunching techniques, a shorter needle would probably be inadequate to reach muscle and might result in increased local reactivity and likely pain and tenderness. Zuckerman therefore suggested the need for individualization of needle length on the basis of patient size and injection technique. The Red Book suggests a needle length of ⅛ inch for newborns to 2 months of age and 1 inch for infants. For toddlers and older children, it suggests ⅝ to 1 inch if the deltoid is used and 1 to 1 ⅛ inch if the anterolateral thigh is used. For adolescents and adults, for whom the deltoid is clearly recommended, the needle length should be 1 to 2 inches. These recommendations are generally in line with the majority of literature reports, and we endorse them.

**Injectate Properties**

**Properties Studied**

A number of investigators have attempted to examine the impact of properties of the injectate itself on the pain associated with intramuscular injection. Three areas have been examined, namely, (1) changing the temperature of the injectate, (2) changing the diluent used to dilute it, and (3) changing the chemical properties of the injectate. Although all of these factors may have a role in reducing injection pain in general, few have been studied with respect to immunization pain specifically.

**Temperature Changes**

It has been theorized that some of the pain of immunization administration stems from the fact that many vaccines are refrigerated just before injection. Because it is generally assumed that injection of a cold substance is more painful than injection of a warm substance, it has been postulated that warming the injectate before administration may reduce pain. Studies on warming of other agents (such as lidocaine) before administration have substantiated this hypothesis. In a study comparing the pain associated with injection of a cold substrate with that of injection of a substrate at room temperature or body temperature, the latter was identified as least painful. With immunizations, however, there is the added complexity of the impact of temperature changes on biological availability and reactogenicity. In fact, in the one study that addressed the issue of injectate temperature and pain in immunizations, the authors found essentially no relationship. Maiden et al randomly assigned patients >16 years of age who required adult diphtheria-tetanus vaccine when they arrived at an emergency department to receive “cold” vaccine, “rubbed” vaccine (rubbed between the palms for 1 minute), or “warmed” vaccine (warmed in a 37°C incubator to body temperature). Although the cold vaccine had a significantly lower temperature when it was administered than did the vaccine in the other groups, there was no difference in time-averaged or peak pain scores. Although this area requires more research, changing the temperature of the immunization cannot be endorsed.

**Type of Diluent**

The choice of diluent may reduce both the short-term and long-term discomfort of an injection. Schichor et al

---

**References**

randomly assigned adolescents scheduled to receive a ceftriaxone injection to 2 groups, one in which the diluent was sterile water and the other in which the diluent was lidocaine. There were dramatic differences between the groups, both initially and at 4 and 24 hours after injection. Amir et al. reported similar results using lidocaine as a diluent for benzathine penicillin G, compared with sterile water. Because a number of vaccines are not premixed (eg, measles-mumps-rubella [MMR], varicella, and Haemophilus influenzae type b) and require dilution, the use of lidocaine instead of sterile water as diluent may be of benefit. Unfortunately, the impact of lidocaine on the biological activity of current vaccines has not been studied; therefore, this practice cannot be supported at this time.

Injectate Formulation
Two studies have compared the associated pain of various formulations of MMR vaccine. In the first study, by Lyons and Howell, the MMR II vaccine (Merck and Co, Whitham Station, NJ) was compared with the pluserix MMR vaccine (GlaxoSmithKline, Middlesex, England) in children whose average age was 44 months. Those given the MMR II vaccine were at least 2 times more likely to cry, compared with those given the Pluserix vaccine. In the more-recent study by Ipp et al., 12-month-old infants were assigned randomly to receive the MMR II vaccine or the Priorix vaccine (GlaxoSmithKline), a vaccine similar to the Pluserix vaccine. Sophisticated pain-assessment measures were used to evaluate the different responses of the infants, including behavior pain ratings based on a review of videotapes of the infants receiving their injections. Visual analog scale scores reported by the parents and the pediatrician were also obtained. There were dramatic differences in pain scores on all measures, with the Priorix group experiencing significantly less pain. Ipp et al. suggested that the pain reduction might be associated with the slightly higher pH of the Priorix vaccine (7.2–7.6 vs 6.2–6.6). The Pluserix vaccine also was less acidic than the MMR II vaccine. This work is in keeping with previous work on the impact of buffering, lidocaine, and subsequent reduction of pain on injection. If available, these vaccines should be used.

DURING THE INJECTION
Parental Behavior
Parental behavior during immunizations significantly influences the amount of pain and distress children experience. (Although the term “parent” is used throughout, most studies involved mothers, not fathers.) For example, Frank et al. found that 53% of the variance in child distress during an immunization was accounted for by maternal behavior. The parental behaviors that were associated with increases in child coping included humor, commands to use coping strategies, and nonprocedural talk to the child. Humor and nonprocedural talk in this context could be considered attempts to distract the child. Empathy, criticism, apologies, giving control to the child, and reassurance were associated with increased child distress. It should be noted that these studies were examining not only immunizations but also other needle procedures, such as bone marrow aspirations and lumbar punctures.

Unfortunately, one of the adult distress-promoting behaviors, reassurance, has also been found to be the most common adult vocalization during immunizations. Although the positive relationship between parental reassurance and child distress is counterintuitive, it has been supported by both correlational and experimental studies. However, not all children may be affected negatively by reassurance. For example, Broome and Endsley showed that some children reacted in a very negative manner (increased distress) to very positive parental behavior (including reassurance), whereas others responded positively.

Although the previously cited work focused primarily on older children, parental impact on the responses to pain in infants also is evident. Sweet et al. found that, at 6 months, maternal behavior accounted for 26% of infant pain behavior, with an additional 17% accounted for by infant difficultness. Furthermore, infant vagal tone interacted with maternal behavior. The pain behavior of infants with high vagal tone was not affected by maternal behavior, whereas maternal behavior did affect the pain behavior of infants with lower vagal tone. Cohen et al. examined discrete parental behaviors during infant immunizations and found that adult distraction was related to infants engaging in distraction and adult reassurance was related to infant distress.

Mothers and fathers may hold differing opinions about reassurance and related constructs. In a study by Schechter et al., fathers were more likely than mothers to agree strongly with the idea that “comforting encourages more crying.” In addition, it may be that parental views of reassurance are changed by personal experience of its negative effect. For example, although parents trained to reassure their children during immunizations were more confident in their ability to help before the procedure than were parents trained to distract, those who reassured were significantly more upset after the procedure.

Although parents can be trained to interact with their children in a manner conducive to less distress and increased child coping, training nurses may affect parental behavior. For example, Cohen et al. found that a distraction coaching intervention was equally effective in increasing child coping and decreasing distress when only the nurses or the parents, children, and nurses were trained. Unfortunately, it is difficult to select parents who need targeted training in effective tech-
niques because parents are not able to give accurate self-reports of their behavior during their children’s immunizations. Cohen et al identified no significant relationship between what parents report they do and what they actually do during an immunization procedure. There was also no significant relationship between what parents say they do and child distress or coping during the procedure.

Nurse and parental behavior may have differing effects on children during immunizations. More specifically, several studies with both infants and older children have demonstrated that nurse behavior is associated with child coping, whereas parental behavior is related to child distress. Sweet and McGrath found that mothers of infants made more coping and distress-promoting verbalizations than did medical staff members across all phases (baseline, injection, and recovery). The impact of adult behavior on infant pain, measured as facial action, varied across phases, with no impact found during baseline. Maternal distress promotion and staff coping promotion accounted for 33% of variability in child pain during the injection phase, whereas maternal distress promotion accounted for 10% in the recovery phase.

Beyond the specific parental behaviors exhibited during pediatric immunizations, research has examined other parental qualities that may affect child behavior, such as parenting style, attachment relationship, and anxiety. Parenting style has also been shown to influence the amount of distress a child demonstrates. Broome and Endsley found that children of authoritative mothers (defined as high in warmth and high in control) were significantly less distressed during an immunization than were children of authoritarian (low in warmth and high in control), permissive (high in warmth and low in control), or unresponsive (low in warmth and low in control) mothers.

In summary, parental behavior during immunization clearly affects the child’s distress during the procedure and the ability to cope. Although it seems counterintuitive, excessive parental reassurance, empathy, and apologies might increase distress, whereas humor and distraction tend to decrease distress. As discussed below, teaching parents techniques that they can use to promote coping should benefit parents by giving them roles that reduce their sense of helplessness and should benefit the child by reducing distress and increasing mastery. Nurse behavior may ameliorate some of the distress-promoting parental behaviors. Although it has not been studied, physician behavior is likely to follow the same pattern.

**Securing the Child**

There is limited literature on how children should be restrained during the injection. For young children, the goal is to have the thigh exposed and relaxed. Often this involves having the parent hold the child in his or her lap. Older children either can sit in their parent’s lap, facing the parent with their legs wrapped around the parent (the so-called “big hug”), or can sit facing forward. The deltoid should be exposed and relaxed if at all possible. Some parents are unable or unwilling to be involved in restraining their child. This preference should be respected if there are adequate office staff members, but the parents should remain present during the procedure, distracting the child during needle insertion and soothing the child after the procedure.

**Distraction**

In the arsenal of nonpharmacologic, pediatric immunization, pain management techniques, distraction is recognized as one of the key interventions. There are a number of possible explanations for the function of distraction. The gate-control theory, one of the early scientific models to explain pain transmission, suggests that the central nervous system modulates the pain experience and thus cognitive attention might affect the processing and perception of pain. Neurophysiologic data support this explanation. For example, the areas of the brain known to process pain stimuli have been shown to be less active during distraction tasks. Consistent with the gate-control theory, the limited-attentional capacity theory suggests that, if some attention is allocated to a distracting task, then there are necessarily fewer resources available to attend to pain. Behaviorally, distraction can be described as introducing a stimulus with previous positive associations (e.g., favorite toy), evoking behaviors that are incompatible with pain behaviors (e.g., laughing at a movie), and reinforcing nonpain reactions.

A number of distraction stimuli have been evaluated for pediatric acute pain. For example, researchers have examined movies, party blowers, nonprocedural talk, interactive robots, virtual reality goggles, kaleidoscopes, bubble-blowing, short stories, and music. In a meta-analysis of 19 pediatric pain management studies using distraction, results suggested a moderate effect size in decreasing distress behavior but minimal impact on self-reported pain. Results also indicated that distraction was most effective for children <7 years of age; there were no gender or ethnic differences in the efficacy of distraction.

Although a variety of distraction stimuli have been examined, few studies have compared them to determine whether one might be better than another. Mason et al compared cartoon movies versus a storybook that required button pressing in an attempt to distinguish passive and interactive distraction with 7 preschool-aged children undergoing Port-a-cath (Smiths Medical, Kent, England) and Hickman line access procedures. Results suggested that the interactive distraction decreased children’s pain more effectively; movies did not differ from
control treatment. The small sample size and the apparent low distraction quality of the selected movies limit the findings. Recently, MacLaren and Cohen\(^95\) found that movies were superior to an interactive robot for venipuncture pain in 88 children. Until more comparison studies are performed, it remains unclear whether some distraction stimuli might be optimal.

Related to the quality of the distraction, MacLaren and Cohen\(^95\) found that, the more children are engaged in distraction, regardless of the type of distraction stimuli, the lower the pain. This inverse relationship between attending to distraction and decreased distress was also found by Cassidy et al\(^96\) in an evaluation of cartoon movies for preschool-aged children’s immunizations. Therefore, the particular distraction stimuli might be less relevant than whether or not they engage the patient’s attention.

The distraction literature is complicated by the fact that many distraction interventions incorporate other psychosocial components, such as role-playing,\(^97\) coping skills training,\(^98\) and positive reinforcement of cooperative behavior.\(^99\) Similarly, the manner in which the distraction is conducted varies across studies. For instance, the clinician performing the procedure has performed the distraction,\(^66\) parents have served as coaches,\(^76\) and the researchers have been the mode of distraction.\(^80\) One of the only studies to compare methods of distraction suggested that using the nurses to coach might be as effective but more practical than training the parents to coach.\(^76\) However, this would likely vary depending on the complexity of the distraction stimulus.

Although distraction for immunization pain has been examined, much of the data on distraction have been obtained with children undergoing other procedures, such as venipunctures\(^99\) or lumbar punctures and bone marrow aspirations.\(^98\) In the immunization pain literature, the majority of work has focused on preschool-aged children, because they exhibit high distress and require immunizations before entering school. In fact, only a few studies have examined distraction for immunization distress in infants.\(^29,100,101\) and one study was performed for immunizations in preadolescents.\(^102\) These studies provide moderate support for adult-led distraction for non-preschool-aged children, but it is clear that more research on distraction for immunization pain in non-preschool-aged children is warranted.

When selecting a distraction stimulus, health care professionals should consider a number of factors. For example, the age and cognitive maturity of the patient is a critical consideration; age-appropriate and engaging stimuli (eg, toys or movies) should be selected. Other factors to consider are the cost, time, and space required for the stimulus. Clinicians might also consider children’s natural coping tendencies\(^95\) and temperament,\(^9\) patient preferences, and other individual characteristics when selecting an intervention. Research suggests that emotional valence and novelty of the distraction stimulus might be important factors, but these variables have not been examined sufficiently to allow direct recommendations.\(^103\) Given the ease of use, the growing body of empirical support, and the lack of apparent negative consequences, clinicians should routinely use distraction for pediatric immunizations. As researchers continue to examine the mediators and moderators of distraction for pain management, clear recommendations might be available to indicate how, when, where, what type of, and for whom distraction is optimal.

**Use of Sucrose**

Sucrose water (12%–50%; typically 1 packet of sugar in 10 mL of water) or other sweet solutions, when administered just before the procedure, have been shown to decrease the pain associated with procedures in neonates.\(^104,105\) It has been suggested that sucrose interacts with opioid pathways to accomplish this phenomenon; however, recent studies call this model into question.\(^106,107\) Sucrose can be administered with a pacifier or directly instilled into the mouth with a small syringe. It seems to work best in the neonatal period and loses its efficacy by 4 to 6 months of age.\(^105\) Nonnutritive sucking (sucking on a pacifier) also seems to have analgesic properties in neonates and may be synergistic with the use of sucrose.\(^106–110\) The combination of direct parental contact and sucrose seems to have an additive effect on the reduction of pain.\(^111\) Breastfeeding has also been shown to provide analgesia when allowed to continue during heel lancing.\(^112\)

Sucrose has been studied for pain reduction associated with immunization. Lewindon et al\(^104\) showed reductions in both crying time and nurse Oucher ratings for infants at 2, 4, and 6 months who received 2 mL of a 75% sucrose solution before immunization, compared with those who received sterile water. Barr et al\(^105\) showed mild reduction of postinjection crying in 2- and 4-month-old infants who received a 50% sucrose solution. Allen et al\(^111\) found 12% sucrose to be no more effective than sterile water but more effective than no intervention in reducing crying times for 2-week-old, 9-month-old, and 18-month-old children receiving a single injection. Neither sucrose nor sterile water provided any benefit, compared with no treatment, for 2-, 6-, and 15-month-old children receiving >1 injection. Crying time was the only outcome studied. Ramenghi et al\(^111\) found that 50% sucrose resulted in shorter crying times, compared with 25% sucrose, glucose, or sterile water, for 4-month-old children. There was a trend toward shorter crying times for 2- and 3-month-old children receiving 50% sucrose in that study, but it did not reach statistical significance. Some studies evaluated sucrose in combination with other techniques. Reis et al\(^111\) found that sucrose in combination with oral stimulation and parental holding decreased crying duration but not.
heart rate changes in 2-month-old children. Lindh et al. showed less crying and less change in modified behavioral pain scores, as well as lower parent and nurse visual analog scale scores, for 3-month-old children who received eutectic mixture of local anesthetics (EMLA) cream plus glucose, compared with placebo cream plus sterile water, before the DPT immunization.

Although sucrose use has been studied and clearly has efficacy for young infants, a systematic research approach to this issue is indicated. This would include testing a variety of sucrose concentrations for 2-, 4-, and 6-month-old children, with and without other modalities such as nonnutritive sucking and parental contact, by using a variety of standard pain measures, including crying time, behavioral pain scores, heart rate variability, and salivary cortisol measurements. At this time, however, there seem to be sufficient data to recommend sucrose use as a routine part of immunization administration for infants ≈6 months of age.

Topical Anesthetics

Topical anesthetics have been shown to reduce the pain of both subcutaneous and intramuscular injections, despite their limited skin penetration. This finding has spurred attempts to develop better topical anesthetics for use in office settings. Such agents should be safe and inexpensive, should have rapid onset (≤10 minutes), and should have no effect on vaccine immunogenicity.

The topical anesthetic most studied for use with immunization is EMLA cream. Despite the limited depth of penetration of EMLA, some studies suggest that it is effective in reducing the pain of both subcutaneous and intramuscular immunizations. The success of EMLA in reducing injection pain is likely attributable to a decrease in pain as the needle penetrates the skin, as well as a reduction in the underlying muscle spasm associated with this pain. EMLA is safe, even for premature infants, and does not alter the immunogenicity of the MMR vaccine.

Topical anesthetics certainly should be considered. Vapocoolant sprays might be used more routinely because they are inexpensive, but their efficacy remains uncertain.

Technical Variables

General Considerations

A number of technical factors associated with the injection itself have been examined, to determine their impact on injection pain. As with other aspects of the injection experience, there are very few studies that help guide practice, despite the frequency with which injections are administered. The studies that exist often have methodologic flaws that limit their usefulness. However, enough evidence exists to allow tentative recommendations for providers.

Injection Technique

There seems to be general agreement regarding the technique of needle insertion. For intramuscular injection, whether in the anterolateral thigh, deltoid, or ventrogluteal area, the general suggestion is that the skin should be held taut around the injection site and the needle darted into the skin at an angle. Traditionally, the plunger is then slightly withdrawn to certify that the needle did not enter a vein. Recently, however, researchers examined the value of this practice and
found it to be unnecessary. The vaccine is then injected under steady pressure, and the needle is withdrawn at the same angle at which it was inserted. If the vaccine is to be administered subcutaneously (inactivated polio, MMR, or varicella vaccine), then the skin around the site is typically pinched or bunched and the needle is inserted at a 45° angle. The rest of the process is identical to that for intramuscular injection.

An alternative technique that was suggested in the nursing literature ~20 years ago is the Z-track technique. In this technique, the injector’s hand is placed on edge across the injection site. The skin is then displaced from the underlying subcutaneous tissue, and the injection is administered. When the injection is completed, the skin is released and returns to its normal position, trapping the injectate in the muscle and eliminating any seepage back into the subcutaneous tissue. In a study by Keen of adults receiving medications other than immunizations, less pain was reported in the Z-track group, compared with the standard-technique group. There are no available data on the frequency with which this technique is now or was ever used.

Site Pressure

It was suggested that pressure at the injection site would reduce the pain associated with immunization. This work is based on the gate-control theory of pain, which posits that flooding an area with a stimulus that is non-noxious reduces the intensity of a painful stimulus in the same general area. The example traditionally offered is rubbing your knee when you bump it on a table.

Barnhill et al randomly assigned 93 adults who were scheduled for intramuscular injection of immunoglobulin to 2 groups; one group received 10 seconds of direct pressure to the injection site, and the other group received the standard injection technique. Those authors found that the pressure group reported a slight but statistically significant reduction in the immediate pain of the injection (13.6 vs 21.5 mm on a 100-mm visual analog scale; \( P = .03 \)).

In a more-recent article, Chung reported on 74 adults who each received 2 injections, 1 with 10 seconds of pressure before the injection and 1 with no previous pressure. Their findings were similar to those of Barnhill et al, although slightly more robust (1.77 vs 2.62 on a 10-point visual analog scale; \( P = .0001 \)).

An extension of this work has been the development of a device known as the ShotBlocker (Bionix, Toledo, OH). The ShotBlocker is a horseshoe-shaped plastic sheet that is tufted on one side. The device is placed around the injection site, and the needle is placed in the middle of the horseshoe. The theory is that the multiple tufts, which are pressed gently against the skin, diffuse the pain sensation. The company that developed this device reported only 2 studies in children with the ShotBlocker, neither of which was published. One study by Guevarra (A. D. Guevarra, MD, unpublished data, 2003) reported the use of the ShotBlocker for 59 kindergarten students receiving an immunization and compared those children with 60 control subjects. Pain was assessed by using the Wong-Baker scales and reportedly 93% of the ShotBlocker group reported no pain, compared with 51% of the non-ShotBlocker group. Similar findings were reported in an unpublished study by Gundrum et al (T. Gundrum, PharmD, C. Sherman, MD, S. Ruhlman, BA, unpublished data, 2001). Both of those studies were methodologically flawed and presented only limited data supporting their conclusions. In contrast, Foster et al reported that the ShotBlocker, in a randomized, controlled trial at an immunization fair, was ineffective in reducing immunization pain. That study was compromised by its broad age range (3 months to 17 years) and the fact that children received 1 to 7 immunizations during the intervention. Although these studies are inconclusive and more rigorous research is necessary, a wealth of anecdotal data support the value of pressure at the injection site as both effective and without adverse effects.

Simultaneous Versus Sequential Administration

Finally, in this era of multiple injections at each visit, the possible advantage of simultaneous injections (multiple staff members injecting at different sites at the same time), as opposed to sequential ones (one after the other), has been suggested. To date, 2 studies have examined this question, 1 in infants and 1 in older children. Those studies yielded essentially the same conclusion; although there was no detectable decrease in discomfort in the child when the immunizations were given simultaneously, compared with sequentially, parents preferred simultaneous administration if possible.

CONCLUSIONS

What emerges strongly from this review is the irony that there is limited research available to address the pain associated with the painful procedure most commonly performed in pediatric office settings. Immunizations have been administered for more than half a century, but the distress they engender has not been examined with the attention or rigor that might be predicted on the basis of their frequency. Perhaps it should not be surprising that procedures that were developed before the burdens of evidence-based medicine continue to be performed without the scrutiny that newer interventions receive. Most of us have long ago made peace with the necessity of intramuscular injections and have developed our own administration strategies that we perceive minimize the associated discomfort. We also know that, for the most part, children tolerate these relatively minor invasions (in fact, they have no choice) with seemingly few long-term ill effects. As a result, there has been no driving force to refine our practices.
When the situation is actually examined, however, there is a price to pay for intramuscular immunizations. As mentioned above, although immunizations are an inconvenience for most children, they are truly dreaded by some. In general, anticipation of an injection fosters an air of persistent tension that hangs over the clinical encounter, especially in this era of multiple injections. This tension is experienced not only by the child but also by his or her parents and the clinician. It behooves us, therefore, to minimize this pain if that can be reasonably accomplished.

This review attempted to collate the relevant literature and to draw some conclusions about the appropriateness and efficacy of currently available interventions. Before any review, it should be noted that, even in areas in which there is general agreement among experts, controversy persists. In addition, the complexities of child development necessitate modifications based on age and understanding.

Before the injection, preparation seems to be beneficial for children >2 years of age. For young children (<4 years of age), preparation should be done in close chronological proximity to the injection itself. For infants and toddlers up to 18 months of age, the current standard site for intramuscular injection is the anterior thigh (vastus lateralis), with a 3/8- to 1-inch needle for children >2 months of age. The ventrogluteal site may be an alternative. At >36 months of age, the deltoid should be used instead of the thigh, but the ventrogluteal site may still be an alternative. Between 18 and 36 months of age, there is controversy regarding the most-appropriate site.

During the injection, there is strong evidence to suggest that parental demeanor has an impact on child behavior. Although it seems counterintuitive, children often are more distressed when parents are more rather than less involved. Therefore, a matter-of-fact, supportive, nonapologetic approach is endorsed. Parents should be instructed to use distraction techniques that are in concert with their personal style and their child’s age and temperament. These techniques may include storytelling, reading to the child, deep breathing, and blowing. Given the expense and time, universal local anesthetic administration is not necessarily appropriate, but use should be considered for children who will require multiple procedures in the future or who are phobic or particularly anxious about the pending injection. Sucrose, which is relatively inexpensive, should be used for children <6 months of age. Pressure at the site, applied either manually or with the aid of a mechanical device, has some support in the literature, is noninvasive and without adverse effects, and is inexpensive. Finally, it is clear that properties of the injectate itself can affect pain, but this area has had almost no sophisticated research and is an open area for investigators to explore. In Table 1, the specific techniques discussed and the levels of support for their efficacy are summarized.

We do not contend that injection pain is a major public health menace. It is clearly an annoyance that affects all children somewhat and some children extremely. It affects their parents and health care providers as well. New research clearly is necessary to better define this area but, as in other aspects of pain management, the systematic application of already available knowledge should go a long way toward reducing pain.

ACKNOWLEDGMENTS

We are indebted to the Mayday Fund for its generous support of the meeting and its long-standing and ongoing commitment to addressing children’s pain. Dr Schechter acknowledges the generous support of the Rockefeller Foundation Bellagio Study Center, where sections of this manuscript were written. We appreciate the invaluable contributions of other members of the consensus group, including Leonard Banco, MD, Bruce Bernstein, PhD, Frederick Bogin, MD, Kathleen McKay, PhD, Lee Pachter, DO, and Harry Weinerman, MD.

REFERENCES


<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Recommended Strategies for Pain Reduction During Immunization and Supporting Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recommendation</td>
</tr>
<tr>
<td>Before injection</td>
<td>Developmentally appropriate preparation for all children &gt;2 y of age</td>
</tr>
<tr>
<td></td>
<td>Needle length, in</td>
</tr>
<tr>
<td></td>
<td>Newborn: 5/8</td>
</tr>
<tr>
<td></td>
<td>Infant: 7/8 to 1</td>
</tr>
<tr>
<td></td>
<td>Older child: 7/8 to 11/4</td>
</tr>
<tr>
<td></td>
<td>Adolescent: 1–2</td>
</tr>
<tr>
<td></td>
<td>Injection site</td>
</tr>
<tr>
<td></td>
<td>0–18 mo: vastus lateralis</td>
</tr>
<tr>
<td></td>
<td>18–36 mo: vastus lateralis/deltoid</td>
</tr>
<tr>
<td></td>
<td>&gt;36 mo: deltoid</td>
</tr>
<tr>
<td>During injection</td>
<td>Parental demeanor should be calm and matter of fact, without excessive reassurance</td>
</tr>
<tr>
<td></td>
<td>Sucrose for children &lt;6 mo of age</td>
</tr>
<tr>
<td></td>
<td>Pressure at injection site</td>
</tr>
<tr>
<td></td>
<td>Use of distraction techniques during injection</td>
</tr>
<tr>
<td></td>
<td>Use of local anesthetics for selected children</td>
</tr>
<tr>
<td></td>
<td>Simultaneous injections if possible, if multiple injections are necessary</td>
</tr>
</tbody>
</table>

The US Preventive Services Task Force levels of evidence are as follows: level A, recommendations are based on good consistent scientific evidence; level B, recommendations are based on limited or inconsistent scientific evidence; level C, recommendations are based primarily on consensus and expert opinion.
15. Reis EC. Multiple scheduled injections contribute to missed opportunities to immunize during well care visits. Ambul Child Health. 1997;3:172
44. Cook IF, Murtagh J. Comparative reactivity and parental acceptability of pertussis vaccines administered into the ventrolateral area and anterolateral thigh in children aged 2, 4, 6, and 18 months. Vaccine. 2003;21:3330–3334
45. Cook IF, Murtagh J. Ventrolateral area: a suitable site for intramuscular vaccination of infants and toddlers. Vaccine. 2006;24:2403–2408
49. Thompson MK. Needling doubts about where to vaccinate. BMJ. 1988;297:779–780
50. MacDonald NE, Marcuse EK. Neurologic injury after vaccination: buttocks as injection site. Can Med Assoc J. 1994;150:326
51. Zuckerman JN. The importance of injecting vaccines into muscle: different patients need different needle sizes. BMJ. 2000;321:1237–1238
54. Davenport JM. A systematic review to ascertain whether the standard needle is more effective than a longer or wider needle in reducing the incidence of local reaction in children receiving primary immunization. J Adv Nurs. 2004;46:66–77
73. Sweet SD, McGrath PJ. Relative importance of mothers’ versus medical staffs’ behavior in the prediction of infant immunization pain behavior. J Pediatr Psychol. 1998;23:249–256


124. Reis EC, Holubkow R. Vapocoolant spray is equally effective as EMLA cream in reducing immunization pain in school aged children. *Pediatrics.* 1997;100(6). Available at: www.pediatrics.org/cgi/content/full/100/6/e5


142. Bogin FJ, Bernstein BA, Payton JS, Schechter NL, Ristau B. A comparison of the pain associated with simultaneous (SIM) vs. sequential (SEQ) immunization injection given at the 9 and 12 month well child visits. *Pediatr Res* 2004;55:210A


Implementing Pay-for-Performance in the Neonatal Intensive Care Unit

Jochen Profit, MD, MPH^a,b, John A. F. Zupancic, MD, ScD^c, Jeffrey B. Gould, MD, MPH^d,e, Laura A. Petersen, MD, MPH^b,f

^a Section of Neonatology, Texas Children’s Hospital, and ^b Section of Health Services Research, Department of Medicine, Baylor College of Medicine, Houston, Texas; ^c Houston Center for Quality of Care and Utilization Studies, Houston Veterans Affairs Medical Center, Houston, Texas; ^d Harvard Neonatal Perinatal Medicine Program, Beth Israel Deaconess Medical Center, Boston, Massachusetts; ^e California Perinatal Quality Care Collaborative, Palo Alto, California; ^f Perinatal Epidemiology and Health Outcomes Research Unit, Division of Neonatology, Stanford University School of Medicine, Lucile Packard Children’s Hospital, Palo Alto, California

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

ABSTRACT

Pay-for-performance initiatives in medicine are proliferating rapidly. Neonatal intensive care is a likely target for these efforts because of the high cost, available databases, and relative strength of evidence for at least some measures of quality. Pay-for-performance may improve patient care but requires valid measurements of quality to ensure that financial incentives truly support superior performance. Given the existing uncertainty with respect to both the effectiveness of pay-for-performance and the state of quality measurement science, experimentation with pay-for-performance initiatives should proceed with caution and in controlled settings. In this article, we describe approaches to measuring quality and implementing pay-for-performance in the NICU setting.
Large deficits in quality of care remain more than half a decade after the Institute of Medicine (IOM) provided a blueprint for improvement. In neonatology, there is persistent unexplained variation in health care delivery and outcomes. To date, quality improvement (QI) efforts, either locally or as part of collaborative efforts, have had mixed results. The broad-based improvement envisioned by health care payers and the IOM has not occurred. (In this article, we use the term “payer” to mean the broad group of employers, purchasers, insurers, and health care plans that pay for health care services directly or indirectly.)

One factor that is receiving increasing attention is a reimbursement system that may actively discourage QI. For example, in December 2003, the New York Times described how Intermountain Health Care, a network of 21 hospitals in Utah and Idaho, was punished financially by Medicare for saving lives and cutting costs. Reimbursement decreased because better care resulted in lower complication rates. In health care, the financial benefits of QI often accrue primarily to payers and patients and not to providers. Pay-for-performance represents an attempt to correct this imbalance and to provide incentives for quality to providers.

By paying providers according to the quality of care they deliver, pay-for-performance schemes attempt to align the interests of health care payers, patients, and providers, ensuring that providers act in the other parties’ best interest. Pay-for-performance initiatives provide financial motivation but may also introduce competitive motivational incentives by comparing the performance of providers again each other or against a standard of care. Pay-for-performance programs thus hold promise for QI by generating both intrinsic (motivation) and extrinsic (reputation and financial rewards) performance incentives.

Although relatively little evidence for their effectiveness has been accumulated to date, 2 comprehensive reviews of the topic found moderate benefits of pay-for-performance and drew cautiously optimistic conclusions about its potential to improve quality of care. In one review, 14 of 17 studies showed partial or positive effects on quality of care. However, it should be noted that, in some studies, improvement owed more to improved documentation than to actual changes in care delivery. Only 3 studies were carried out in the pediatric population, and all targeted preventive care services in the general pediatric health care delivery setting.

Despite some ambiguity in early evaluations, the IOM has endorsed ongoing experimentation with pay-for-performance, and payers are enthusiastic about its potential to improve the value of health care purchasing. There are now >100 active pay-for-performance projects throughout the country. In addition, legislative initiatives aim to incorporate incentives for quality into Medicare’s payment systems. Although to our knowledge pay-for-performance approaches have not been applied in the NICU, we think that the NICU is a prime target for payers because of the high cost, available databases, relative strength of research evidence, and, compared with adult settings, low incidence of comorbidities. The latter makes it easier to attribute performance to providers, rather than to patients.

Unfortunately, many pay-for-performance projects are implemented in an uncontrolled manner, making it unclear whether the benefits are truly attributable to the financial incentives. Rigorous research designs and methods are necessary to determine whether performance-based payment arrangements result in meaningful QI and are cost-effective. For example, 2 of us (Drs Petersen and Profit) are conducting a prospective, multicenter, cluster-randomized, controlled trial to study the effects of the pay-for-performance approach on quality of care and hypertension control in adults (L.A.P., L. D. Woodard, MD, T. Urech, MPH, et al, unpublished data, 2007). That trial should add to the body of literature on pay-for-performance and shed light on the benefits and costs of different choices in incentive design. It uses physician- and group-level financial incentives, plus audit and feedback, to improve quality of care. More such trials need to be designed to evaluate the effectiveness of pay-for-performance in a variety of care settings and for a spectrum of clinical situations. Our recommendations for implementing quality assessment and financial incentives for future pay-for-performance initiatives in neonatology are described below and summarized in Table 1.

**MEASURING QUALITY**

**General Approach**

Defining and measuring quality is central to a pay-for-performance program’s relevance and ability to meet its objectives. Careful attention to quality measurement is also important for the feasibility of implementation, because physician opposition to pay-for-performance in health care is often grounded in questions about the basic validity, fairness, and meaningfulness of the assessment methods. A fair and scientifically sound approach to quality measurement may enhance provider accep-
tance and alleviate concerns that pay-for-performance is primarily a cost-cutting measure, rather than a QI tool. Although policymakers and payers are moving ahead with pay-for-performance, the science of quality measurement has not kept pace, which has created a serious disconnection between policy intentions and potential outcomes. Measures that define clinical quality too narrowly or lack the support of empirical evidence are unlikely to yield the desired improvements in health care quality. Furthermore, the process of measurement should minimize undesirable secondary effects on physician behavior. For example, when pay-for-performance initiatives provide financial incentives for quality, measures that are not adjusted for clinical risk offer an incentive for providers to avoid treating the sickest patients.

**Framework for Measuring Quality**

Generally, quality of care is defined within a multidimensional framework. For example, the IOM has suggested that quality of care is a reflection of care in the domains of patient safety, effectiveness, efficiency, patient-centeredness, timeliness, and equity. The dimensions of the quality of health care delivered by a NICU may also be described by its physical and organizational composition (structure of care), by the clinical care interactions between patients and providers (process of care), and by patient outcomes, in terms of morbidity, death, and caregiver satisfaction (outcomes of care). Measures of structure, process, and outcome have distinct advantages and disadvantages. For example, structural measures (eg, the availability of electronic health records) are easy to obtain and measure but are theoretically distant from the ultimate goal of improving health outcomes. Process measures may be more sensitive to differences in quality of care but require that there be good evidence for a direct link between the process and clinical outcomes. Outcome measures are perhaps of greatest intrinsic value, because they reflect directly what patients and providers truly care about, but they may occur too infrequently to provide statistically meaningful results (eg, death) or may occur so far in the future (eg, developmental delay) that data collection efforts become impractical or burdensome.

Ideally, we think that an assessment of quality should incorporate the full range of quality-of-care dimensions, with indicators that are valid, reliable, feasible to collect, and relevant to important domains of care. Quality assessment is a dynamic process and, especially within pay-for-performance schemes, should reinforce providers’ control over their performance. Accordingly, indicators should be not only theoretically sound but also actionable; that is, indicators should be responsive to change within a timely period and should be unambiguous with respect to interpretation. Importantly, measures must be standardized and adjusted for clinical risk, and data collection must be adequately simplified to ensure uniformity of definitions.

Figure 1 presents a proposed framework for neonatal

---

**Figure 1**

Theoretical framework for measuring the quality of neonatal intensive care. Solid arrows indicate interactions, and dashed arrows indicate the potential use of a composite indicator to measure health care delivery, to predict health status, and to guide health policy-making at the health systems and societal levels.
quality measurement. Pay-for-performance programs attempt to measure and reward the quality of the products of the health care delivery system. The outcomes of the health care delivery system are influenced by individual and societal determinants of health, as well as the design of the health system. The combination of the structure/process/outcome framework of quality with that of the IOM results in a quality-of-care matrix that forms an inclusive framework for measuring quality. In our opinion, this could address some of the shortfalls of focusing on individual measures. Although identifying the specific indicators for each of these domains of quality might prove challenging, this framework provides a guide to practitioners and researchers in an ongoing effort to refine quality measurement. Evidence-based expert consensus could be used to fill the matrix and to generate measures for quality-monitoring or pay-for-performance initiatives.

**Composite Indicators**

To benchmark and to reward NICUs for high quality of care that has been assessed across several dimensions, the individual dimensional measures of quality need to be combined and aggregated into a summary measure or composite indicator of quality. Such a composite indicator reflects judgments regarding the relative importance of each measure. In other fields of medicine, composite indicators (or scorecards) that capture multiple dimensions of quality have improved the quality of health care institutions, providers, and patient care. A composite indicator of NICU quality could offer performance targets for improvement by showing the gaps between NICU performance and benchmarks. It could provide a foundation for the development of public and private policy action and a yardstick against which to measure the success of new policies. Although the literature describes many different guidelines and methods for constructing composite indicators, a particularly explicit methodical approach has been described by the Organisation of Economic Cooperation and Development (OECD) and applied to several indicators of country performance. Crucially, the OECD guidelines ensure transparency of the composite indicator-building process and promote internal and external statistical and methodologic consistency, an improvement over many currently applied methods.

**OECD Guidelines for Constructing Composite Indicators**

Briefly, the OECD suggests a 10-step building process (Table 2). At each step, researchers must choose from several available options, depending on the underlying data and the purpose of the composite indicator.

**Table 2 Developing a Composite Indicator**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Development of a theoretical framework</td>
</tr>
<tr>
<td>2</td>
<td>Measure selection</td>
</tr>
<tr>
<td>3</td>
<td>Initial data analysis</td>
</tr>
<tr>
<td>4</td>
<td>Imputation of missing data</td>
</tr>
<tr>
<td>5</td>
<td>Normalization</td>
</tr>
<tr>
<td>6</td>
<td>Weighting and aggregation</td>
</tr>
<tr>
<td>7</td>
<td>Uncertainty and sensitivity analysis</td>
</tr>
<tr>
<td>8</td>
<td>Linkage to other variables</td>
</tr>
<tr>
<td>9</td>
<td>Deconstruction of the composite indicator</td>
</tr>
<tr>
<td>10</td>
<td>Presentation and dissemination</td>
</tr>
</tbody>
</table>

Step 2 is measure selection. Importance, accuracy, and feasibility guide the selection of quality-of-care indicators. The medical literature and expert opinion can provide guidance.

Step 3 is initial data analysis. The underlying nature of the data must be explored and appropriate transformations made with regard to directionality of measures, outliers, ceiling effects, and nature of distributions.

Step 4 is imputation of missing data. The impact of missing data on the performance measurement must be examined, because the data may contain significant bias if providers avoid reporting poor outcomes.

Step 5 is normalization of data. For linkage of measures, the measures must be transformed into a common unit of measurement. There are many options for normalization, including ranking, standardization, and distance to a reference.

Step 6 is weighting and aggregation. This is a crucial step in the development of a composite indicator, because the attribution of weights to different measures and their aggregation can have significant influences on performance. The 2 basic approaches used to arrive at subindicator weights include statistical (eg, principal-component analysis, factor analysis, multivariate techniques, and others) and participatory (variations on elicitation of expert opinion) methods. It is important to realize that equal weighting does not imply an absence of weights, because with this approach each subindicator is given a weight of 1. The benefit of the statistical approach includes its relative fairness and freedom from bias in deriving weights based on purely statistical grounds. Its disadvantage is that the weights may not correspond to real-world common sense.

In the aggregation phase, the subindicators are aggregated into a composite indicator. The primary decision involved in choosing an aggregation method is whether NICUs should be allowed to compensate for poor performance in one subindicator with superior performance in others. There are 3 principal choices, namely, full compensation (linear additive aggregation), partial compensation (geometric or multiplicative aggregation), and no compensation (noncompensatory methods). Each of these choices has benefits and drawbacks.
Step 7 is uncertainty and sensitivity analysis. There are 2 primary sources of error in performance measurement, that is, the effect of the error contained within the underlying data (uncertainty analysis) and the impact of different choices in constructing the composite indicator (sensitivity analysis). These error sources can be combined and their effect displayed in a higher-order Monte Carlo experiment.

Step 8 is linkage to other variables. Composite indicators for some fields of medicine might be combined with those in others, potentially yielding greater insights across care settings or longitudinally. Entire networks of care could be compared with respect to their performance in managing acute and chronic care (ie, combining NICU care with follow-up care).

Step 9 is deconstruction of the composite indicator. Both summary scores and performance on individual measures can be displayed to guide health policy-making and future research. This allows stakeholders to identify areas of weakness and strengths.

Step 10 is presentation and dissemination. Results can be presented in user-friendly formats such as charts that include measures of uncertainty (confidence intervals). Electronic publications can link to additional details on individual subindicators.

Measuring Quality in the NICU Setting

Data collection efforts in neonatology are better developed than in many clinical specialties. The Vermont Oxford Network collects validated data from >600 NICUs throughout the world.50 In California, 120 NICUs submit an expanded data set, with core elements identical to those collected by the Vermont Oxford Network, to the California Perinatal Quality Care Collaborative (CPQCC). These data are used to prepare confidential reports for each NICU and to prepare the California Children’s Services mandated yearly activity and outcomes report, which CPQCC submits on behalf of requesting NICUs. A quality indicator based on routinely collected data could thus be used for comparative benchmarking efforts involving pay-for-performance programs. We are currently working to develop such an indicator by using the CPQCC database. A possible representation of NICU quality measures within the matrix is given in Fig 2.

A danger of tying payments to performance is that data quality may suffer as providers use undesirable behaviors, such as omitting poor outcomes from their reports, in order to improve their ratings without improving actual performance. Minimizing such behaviors will require costly, ongoing validation of randomly sampled data, a significant disadvantage of pay-for-performance initiatives. Several other challenges to measuring quality are particularly prominent in the NICU setting and require special attention.

The first challenge involves the diversity of populations. Pathologic conditions, care practices, and outcomes vary widely for patients in different gestational age groups, requiring in some instances both stratum-specific analyses and individualized quality-of-care measures for specific subpopulations, such as extremely premature infants, infants requiring complex surgery, and infants with congenital anomalies. Rather than attempting to measure care for all groups at once, stakeholders should focus on developing quality measures for patient groups that are commonly represented in NICUs (very low birth weight infants, moderately premature infants, and term infants).

The second challenge involves the limit of viability. There is no consensus regarding the treatment of patients born at gestational ages of <25 weeks.51 This group of patients may require a special set of quality markers that relate more to patient satisfaction with care or documentation of parental education than patient-specific outcome measures.

The third challenge involves patient transfers. It is currently difficult to track patients’ hospital stays across multiple institutions of care. This may induce significant bias, because NICUs might transfer their highest-risk patients to other hospitals.52 Another source of bias stems from the differing availability of back-transport across NICUs. Lengths of stays are increased in NICUs where opportunities for back-transport are limited. Evaluations of quality therefore need to account for transfer bias. Risk adjustment should also account for the location of birth (inborn/outborn). Ultimately, improvements in patient tracking may eliminate this problem.

![FIGURE 2](https://www.pediatrics.org)  
Quality matrix filled with measures from the CPQCC database. NICI indicates nosocomial infections; CLD, chronic lung disease; LOS, length of stay; ROP, retinopathy of prematurity; IVH, intraventricular hemorrhage; PVL, periventricular leukomalacia. Only severe degrees of retinopathy of prematurity and intraventricular hemorrhage would be included.
DESIGN OF THE FINANCIAL INCENTIVE

Overall Design
Designing financial incentives is a complex process involving decisions about the structure of the incentive (competitive or noncompetitive), the recipient of the incentive, the amount of the incentive, the structure of the payment, and the frequency of payment. Choices in any of these categories have advantages and disadvantages and must reflect an optimal balance between the incentive’s aims and practicality constraints.

Incentive Structure
Incentive structure influences how rewards are allocated across providers, whether providers compete for bonuses, and whether targets are based on improvement or just good performance. Competitive bonus programs provide an incentive to improve performance as providers compete for rewards and reputation. However, most of the payouts go to the top-performing providers, with little incentive for bottom-performing providers to improve. In noncompetitive programs, all providers are rewarded for reaching fixed performance targets. Targets based on QI rather than absolute quality provide greater incentives for those with low baseline quality, although most of the payouts again go to the high performers. Our preferred approach would be a combination of methods in which providers are rewarded for achieving the desired result in any given measure of care but also are rewarded for overall performance and/or improvement on a composite measure of care.

Incentive Recipient
The more direct the connection between the incentive and the person delivering the care, the greater is the effect of the incentive. In the NICU setting, however, care practices and results rarely can be attributed to a single provider but rather are a reflection of a team effort that includes a group of caregivers (eg, physicians, nurses, respiratory therapists, and nutritionists). In addition, some patients require multidisciplinary care from surgeons, cardiologists, and other providers. Therefore, in the NICU, a group or hospital incentive is a more-practical design choice. Any financial reward to providers would be redistributed within the group. This design would also foster a collaborative approach to patient care, because all caregivers would participate in the benefits of the reward, although a potential problem with this approach is “free-riding” by providers who contribute relatively little to care improvement within the group.

Incentive Amount
The amount of money needed to change provider effort is variable and is determined by the provider’s marginal utility for the extra income. This depends not only on monetary factors (household income) but also on nonmonetary factors (personal ethics, normative professional practices, regulatory control, and clinical uncertainty). An amount too small is unlikely to induce a change in behavior; an amount too large may induce undesirable provider behavior. A survey of health maintenance organization managers indicated that a bonus of at least 5% of a physician’s capitation income would be required to influence provider behavior.

Payment Structure
The principle choice is whether to reward providers through an intermittent bonus or an increase in the fee-for-service schedule. Economic theory suggests that providers would respond most to incentives if they are rewarded every time they do the right thing or achieve a desirable outcome. However, the psychological literature suggests that larger intermittent bonuses for achieving a benchmark of care may create a more powerful motivational effect than regular small payment increases. There is insufficient literature to make a definitive judgment with regard to either method. For practical reasons related to data collection, we recommend a yearly bonus.

Payment Frequency
Practical impediments to rewarding providers with frequent timely payments to sustain momentum for improvement include the need to collect and to evaluate data. In addition, the frequency depends on the interval of measurement that allows for a meaningful interpretation of change. Specifically, if a measured variable occurs relatively infrequently, then it will take a longer time before a true performance assessment can be obtained. In the NICU setting, at a minimum, yearly feedback would be desirable.

BRINGING PAY-FOR-PERFORMANCE TO THE NICU
There is enormous political momentum from health care payers to realign the payment system to reward providers for the quality rather than the quantity of care. However, because the evidence base on pay-for-performance is still in its infancy, we recommend a phased approach to implementation in the NICU setting. Such an approach would use limited funding in select NICUs for specific quality measures. Systems established on the basis of “best guesses” must be designed with the potential for later revision as data are collected. Throughout the process, it is crucial that providers be involved in decisions about defining and measuring neonatal intensive care quality. In addition, controlled trials need to investigate the benefits and potential pitfalls of pay-for-performance in the NICU setting. Heightened attention to measuring (and rewarding) NICU quality should provide unprecedented opportunities to develop valid...
methods for assessing quality of care, which could have both economic and scientific implications.

REFERENCES


IN THE WORLD OF LIFE-SAVING DRUGS, A GROWING EPIDEMIC OF DEADLY FAKE

“Asia is seeing an ‘epidemic of counterfeits’ of life-saving drugs, experts say, and the problem is spreading. Malaria medicines have been particularly hard hit; in a recent sampling in Southeast Asia, 53 percent of the anti-malarials bought were fakes. Bogus antibiotics, tuberculosis drugs, AIDS drugs and even meningitis vaccines have also been found. Estimates of the deaths caused by fakes run from tens of thousands a year to 200 000 or more. The World Health Organization has estimated that a fifth of the one million annual deaths from malaria would be prevented if all medicines for it were genuine and taken properly. ‘The impact on people’s lives behind these figures is devastating,’ said Dr Howard A. Zucker, the organization’s chief of health technology and pharmaceuticals. Internationally, a prime target of counterfeiters now is artemisinin, the newest miracle cure for malaria, said Dr Paul N. Newton of Oxford University’s Center for Tropical Medicine in Vientiane, Laos. His team, which found that more than half the malaria drugs it bought in Southeast Asia were counterfeit, discovered 12 fakes being sold as artemunate pills made by Guilin Pharma of China.”


Noted by: JFL, MD
Healing of Hymenal Injuries: Implications for Child Health Care Professionals

Howard Dubowitz, MD, MS

Department of Pediatrics, University of Maryland School of Medicine, Baltimore, Maryland

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

The article by McCann et al1 in this issue of Pediatrics provides additional information on the natural course of hymenal injuries in both prepubertal and pubertal girls. In general, they heal rapidly and often completely, confirming similar findings in earlier studies.2,3 This new study raises several implications for pediatricians and other health care professionals who examine children and youth.

It is noteworthy that it took this major US (and Australian) effort involving 25 medical centers to gather retrospective information on 239 girls, both prepubertal and postpubertal, with hymenal trauma. This largely reflects how unusual these injuries are; relatively few girls evaluated for sexual abuse have abnormal findings on physical examination.4-6 Fortunately, most sexually abused girls do not experience serious physical trauma, primarily because of the nature of the abuse. The low base rate of physical injuries, however, has been a challenge for research in this area. Even in this study of 239 girls, there were few girls with certain injuries (eg, 7 prepubertal girls with “transsection only”). Small numbers always demand cautious interpretation. It is noteworthy that none of the girls in this large series had any hymenal scarring, as reported by others.2,7 However, these other authors did note occasional scarring of adjacent structures (eg, the posterior fourchette); the McCann et al article is limited to hymenal findings.

Previous studies have reported that hymenal trauma mostly occurs posteriorly.7 McCann et al1 document hymenal lacerations occurring anteriorly in 5% of prepubertal and 15% of pubertal girls. It seems that none of those girls also had posterior injuries. Thus, McCann et al recommend that medical evaluations for sexual abuse in girls should always include both the prone knee-chest and supine positions, because some anterior and lateral lacerations were found only in the prone position. This finding may be explained by their method of reviewing photodocumentation rather than performing actual examinations. It is probable that the anterior aspect of the hymen may not have been readily visualized in images taken in the supine position because of the positioning of the camera. Examining girls in the prone position is thought to be potentially embarrassing; this position is often only used when the examination in the supine position raises concerns. Acute trauma is invariably visible in the supine position, making the prone examination usually unnecessary. Beyond the “acute” period, many nonabused girls normally have clefts and areas of “missing” hymen anteriorly. This makes the prone examination redundant unless an additional view is necessary to clarify uncertain findings in the supine position.

A goal of the article was to assist with the timing of the hymenal injuries. The retrospective design, with varying periods to follow-up, made this difficult. The authors found, however, varying patterns of healing that suggest timing these injuries was difficult and of limited forensic value. Of note, they do report that petechiae had all disappeared within 2 to 3 days and that blood blisters lasted up to 1 month. These findings may help corroborate histories of when the alleged abuse occurred.
It has long been thought that estrogen-induced changes of the hymen usually mask previous evidence of trauma. The McCann et al article highlights the potential for complete healing in prepubertal girls. In addition to the customary caution that a “normal” examination does not rule out sexual abuse, there is the need to acknowledge that current hymenal normalcy does not preclude earlier injury. Even many of the more serious injuries healed without leaving a trace. Their study confirms the need for considerable constraint in interpreting the physical findings (ie, likelihood of abuse) beyond the acute period.

The medical evaluation is one component of the comprehensive assessment of children suspected of having been sexually abused. Questions remain as to where, when, and by whom these evaluations should be performed. Pediatricians might even question whether the medical piece is really necessary given that so few children are injured and have medical evidence of abuse. First, even if only a few have findings, the medical evidence can make a significant difference in helping protect a child, identifying those responsible, and protecting other children. In addition, some, particularly adolescents, have sexually transmitted infections that need to be treated. Perhaps most importantly, the medical evaluation can provide immense comfort to a child and parents that they are normal, without any residual defect. Children and parents may have an array of fears about being “damaged,” infections, future sexuality, and childbearing. Allaying such anxiety is very important; parents have reported much satisfaction with interdisciplinary evaluations for sexual abuse.

When and where should the medical evaluations be conducted? McCann et al have demonstrated that the physical evidence may disappear within a few days, especially petechiae, the most common finding. Christian et al also reported that forensic evidence, including blood and semen, is mostly found within the first 24 hours, a finding that was confirmed recently by Young et al. Thus, the medical evaluation should be conducted soon after the alleged abuse occurred. Children often delay reporting their abuse, which makes this difficult: once several days or weeks have elapsed, the likelihood of finding clear physical evidence is very slim. Ideally, all communities should have resources for a skilled medical evaluation to be made available quickly (within a day of the alleged abuse). At the same time, it does not mean that these children must be examined immediately unless there is a compelling medical indication (eg, active bleeding, suicidal ideation). Too often, the perceived need for an immediate evaluation leads to an emergency department, in the middle of the night, with staff not trained to evaluate these children and in a setting that may exacerbate their stress.

The concern of acute or recent sexual abuse raises legal issues and the possible need to gather forensic evidence. Few primary care pediatricians are likely to feel comfortable in this role; time demands are an added barrier. A few pediatricians, however, do set aside regular time to provide this service in their offices. North Carolina, for example, has a network of primary care physicians who perform these evaluations on a fee-for-service basis. Within the United States, approximately 600 child advocacy centers have been developed to comprehensively evaluate possible sexual abuse, although many do not have a medical component. Pediatricians can play a valuable role in collaborating with other professionals in these centers. Currently, however, there is a shortage of physicians who are trained and willing to conduct these evaluations, which has contributed to forensic (or sexual assault) nurse examiners often filling this gap, usually in emergency departments. Nurses started these programs to serve adult rape victims, and some of them have received special training to evaluate children. There are a variety of models. Some nurse examiners are supervised by pediatricians; they conduct the evaluations but do not interpret the findings or render a definitive opinion as to the likelihood of abuse. Others function quite independently. There is debate among pediatricians as to the appropriateness of nurses in these roles; research is needed to evaluate these models of care. At least until there are enough physicians able and willing to evaluate these children, nurses will likely continue to provide this service.

The McCann et al study is interesting, because it indirectly shows where earlier thinking turned out to be erroneous. For example, until recently, the diameters of the hymenal opening were examined and reported. McCann et al appropriately do not even report those data. This is important for pediatricians to be aware of, especially when faced with a presenting complaint such as “she looks so big down there.” There may still, however, be a need to probe the concern of possible sexual abuse. Their article does describe thinning (<1 mm) of the hymen in some girls. It is difficult to measure this thinning; hence, there is questionable significance to this finding.

It is the nature of science that knowledge evolves. However, there is an ethical concern here if such previous knowledge contributed to a finding of sexual abuse, a child being removed from the home, and someone being incarcerated. In general, these decisions were based on more than 1 physical finding. Nevertheless, there may be a need to “set the record straight” despite the logistic challenges of whom to contact, how to reach them, and over what period of time.

The McCann et al article also illustrates the importance of precise terminology to help us speak a common language. The authors offer useful guidance by carefully defining their terms. Nevertheless, even with their glossary, some terms are used interchangeably. For ex-
ample, why have both “hymen” and “hymenal membrane” or “complete tear (or laceration)” and “transection?” “Cleft” and “notch” are also very similar terms. Having just one term for each finding would facilitate communication for medical, legal, and research purposes.

The article by McCann et al reinforces a lesson we have learned: the child’s history is the most important aspect of evaluating possible sexual abuse. Physical findings are few and, once the wounds heal, are often subtle and ambiguous or entirely gone. Girls who have been sexually abused generally have the same physical appearance as girls who have not been abused. Thus, although these children need a skilled and early medical evaluation, it is especially important that a detailed history be obtained. The physician’s history can complement the forensic interview conducted by law enforcement or child protective services. The optimal evaluation of children suspected of having been sexually abused demands interdisciplinary collaboration, which should include mental health colleagues for providing treatment.

REFERENCES

Women in Pediatrics: Recommendations for the Future

Women Chairs of the Association of Medical School Pediatric Department Chairs

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.

Task Force on Women in Pediatrics

The Federation of Pediatric Organizations (FOPO) recently issued the “Report of the Task Force on Women in Pediatrics,” and the Association of Medical School Pediatric Department Chairs (AMSPDC) has endorsed the recommendations made. The current women chairs of the AMSPDC support the goals set forth in the FOPO report and believe that successful realization of the FOPO goals requires an understanding of the changing demographics and expectations of physicians who choose careers in academic medicine along with an appreciation of the importance of work/home balance in the successful career development of academic physicians. Here we identify 4 specific areas of focus and provide suggestions for implementation that are consistent with the FOPO report. This commentary is directed to the senior leadership of medical schools and academic medical centers who share an interest in developing effective strategies and systems that will enhance the career development of women in academic medicine.

BACKGROUND

Over the past few years, ~50% of medical students have been women. The specialty of pediatrics has experienced a steady increase in the percent of women entering the field. In 2002, women comprised nearly 70% of pediatricians in training and 50% of all practicing pediatricians. Although the percent of women in medicine continues to represent only 14% of tenured faculty and 12% of full professors. Although the statistics for departments of pediatrics are somewhat better, with 19% of women at the full professor rank, only 14 (9.5%) of the 147 current members of the AMSPDC are women.

There is substantial evidence that women have not advanced in academic rank as rapidly as men in medical schools and are less likely to be considered for or hold leadership positions in academic medicine. A traditional explanation for these findings has been that the pool of qualified women to be considered for such positions is too small because of an inadequate pipeline. We can no longer content ourselves with this “pipeline” theory; the pipe has been primed for well over a decade, and the results have been discouraging. This commentary is not only a call for equity but also a call to improve the quality of our profession by engaging and nurturing the best talent available, including both women and men. Current trends indicate that the profession of medicine will be composed of equal numbers of men and women over the next generation. If we do not change the system to encourage and enable women to contribute maximally to their profession, academic medicine will lose a major source of potential talent.

We have chosen to focus on 4 issues that greatly affect
women in academic careers: (1) the option to work part-time at specific career stages; (2) the availability of high-quality child care; (3) the need for flexibility in the career paths of physician-scientists; and (4) the desire to draw more women into leadership positions in academia. It is our intent to support a new paradigm that encourages productivity among women by allowing them to design their career paths to be more family-friendly rather than trying to mold their career to a more traditional model that was developed decades before the current predominance of the 2-career family. Although we often talk about “work/life balance,” we need a system that not only allows but also encourages different, sometimes nontraditional, career trajectories. Although this commentary is written from the perspective of female chairs of pediatrics with the intent to encourage more women to enter academic fields and aspire to leadership positions, many of the concepts discussed will support career advancement and promote job satisfaction among men as well as women. For example, many men in generation X are also seeking work environments that will allow them to have a high-quality personal life outside the hospital and/or office. Thus, many of the recommendations offered here apply to the professional development of both women and men.

The following 4 areas need systematic thinking, explicit policy development, and strategic reorganization throughout academic medicine.

**THE OPTION TO WORK PART-TIME**

Women often have multiple roles beyond work, including responsibilities to their families. Although there are federal mandates regarding family and medical leave, women often need flexible work schedules or part-time employment to meet concurrent ongoing responsibilities outside of the workplace. The optimal time for childbearing and child rearing occurs at the same traditional time for career advancement in academic medicine, which means that many parents are in the untenable position of having to choose between the needs of their families and the demands of their profession. Having the option of working part-time, at least for a few years, can be a beneficial situation for parents and employers. Although part-time work often incurs extra costs for an institution, the positive impact it may have on faculty morale and institutional loyalty should offset any short-term financial burden. In addition, significant benefits may also be realized with faculty retention and recruitment.

Pediatric departments have more part-time faculty members than other medical school departments, and the overwhelming majority of part-time faculty members are women. In many institutions, these individuals feel marginalized from mainstream academic activities. We believe that department chairs and division chiefs should make an effort to appropriately integrate part-time faculty members into the administrative functions and activities of their unit and department and ensure that they are not marginalized as a consequence of their part-time position. This effort should include consideration for committee assignments and leadership/administrative roles. We would also counsel part-time faculty members to be receptive to invitations to participate in various department activities as one way to stay engaged in the department. Over time, some women may return to full-time positions as their children grow older.

We realize that part-time positions may not always be available in all departments. We also recognize that the needs of full-time faculty members (both men and women) must be considered and supported so that feelings of inequity do not arise and every faculty member believes that he or she is being treated fairly. This may require some creativity on the part of departmental leadership.

Although requests for part-time status most often occur when women have young children, many faculty members, like the rest of society, experience the stress of the “sandwich generation” as they struggle to care for their own children as well as their aging parents. In many instances, it may be appropriate to work part-time while caring for elder relatives. Like infant care, eldercare usually falls on the woman’s shoulders in most families.

Therefore, we encourage all department chairs to seek means to allow faculty members to work part-time because of family responsibilities if requested and to find creative ways to engage such faculty members in academic and departmental activities.

**THE NECESSITY FOR AVAILABLE HIGH-QUALITY CHILD CARE**

Although substantial numbers of women have chosen careers in medicine for more than 25 years, less than half of academic medical centers have on-site day care. Furthermore, almost no centers have day care facilities that are sufficient to meet the demand. Having safe, high-quality child care is the single major factor that young physician-parents must consider in accepting any professional position. Concerns about child care are major impediments to women who work long hours and may significantly interfere with their productivity at the workplace. The ability of women to work in close proximity to their infants and young children facilitates breastfeeding, allows women to feel involved with their child’s care, and reduces the emotional tension often felt by young working mothers. In addition, on-site day care allows young female professionals to socialize with each other. Social isolation is a commonly expressed concern of professional women with children. They feel isolated from colleagues at work because of fears that they are perceived as not “pulling their own weight,” especially if they choose to work part-time or need to leave early or
be absent because of child care issues. At the same time, they may feel isolated from young women who do not work or from women whose work hours are more circumscribed because of choices they have made regarding professional commitments.

Although it is the responsibility of departments of pediatrics and academics in general to be family-friendly and flexible, it is not within their purview to provide day care. We would argue, as has industry, that on-site day care results in stronger family attachment and more productive employees. In this context, it may be true that similar benefits would be realized in academic medical centers.

On-site day care is one model for child care, but this model may not meet the needs of all medical parents, because their professional responsibilities extend well beyond the usual hours of commercial day care whether provided on-site or off-site. These families may require in-home child care (eg, nannies, au pairs, etc) or child care assistance for an unexpected situation in which a child or baby-sitter is sick. To assist those parents, several institutions have contracts with emergency child-sitting services that fill in temporarily (eg, Parents-in-a-Pinch; Caregivers on Call). Usually, parents pay a reduced fee for such services.

Therefore, we encourage all departments of pediatrics to assume a leadership role in their institutions to develop day care programs that are adequate in quality and quantity to serve the needs of their institutions. Furthermore, we encourage academic medical centers to explore creative models to provide child care assistance for medical faculty and staff.

THE NEED FOR FLEXIBILITY FOR FEMALE PHYSICIAN-SCIENTISTS

Fewer women than men choose the physician-scientist career path. There may be several reasons for this. After completing residency or fellowship, young faculty may feel confident of their clinical skills but have less experience and confidence in their research abilities. In addition, the physician-scientist path is a tenure-track position in most universities, with the traditional restrictions and requirements to attain tenure. There is usually a restricted time frame in which one must achieve tenure coupled with explicit requirements for obtaining grant funding and producing publications. In most institutions, part-time work in such a track is not encouraged; the tenure clock is not altered but for rare exceptions; and, for women, the tick of the tenure clock coincides with the tick of their biological clock to bear children. These temporal demands coupled with the need to assume major responsibilities for child care often lead talented women to forego a career as a physician-scientist, which is a loss for our profession as well as for that particular woman. We offer the following suggestions:

1. Tenure-track appointments must have some flexibility. In many cases, the deterrent to implementing this recommendation is university policy, over which many medical school deans and chairs have limited influence. Some universities permit faculty members to stop the tenure clock with a reduction in full-time appointments to 75% to 80%. Often, these appointments allow faculty members the opportunity to maintain funding through foundation and National Institutes of Health granting agencies. Unfortunately, other universities permit only those who work full-time to remain on the tenure track. The security of knowing that a successful academic career can be pursued at less than full-time status may encourage young women to choose an academic career at the early stages of medical training. Even at institutions that grant tenure to faculty who work part-time, there is increasing evidence that those who work part-time are more likely to choose clinical rather than academic/tenure tracks. To attract women to the physician-scientist role, medical school leaders must be willing to mount a cogent argument to have outmoded policies changed.

2. Research/funding institutes, including but not limited to the National Institutes of Health, must allow part-time faculty to apply for grants. The perception that recipients of K awards are not permitted to work part-time has often been cited as a rationale for the requirement for full-time employment as a condition for tenure-track status and has clearly been a deterrent to some women who might otherwise pursue research during their childbearing years. Such permission should be the norm, without requiring special petition. There is considerable evidence that women are concerned that requesting “special” permission places them at career risk.

3. For institutions that retain time-limited tenure decisions, serious consideration should be given to lengthening the tenure clock for women and men who are primary caregivers. In an informal survey of AMSPDC members, 50% of responding pediatric chairs stated that the tenure clock could be stopped for varying periods at their respective institutions, and 40% said that the time to tenure could be lengthened. The potential for stopping the tenure clock is a component of the 2001 American Association of University Professors “Statement of Principles on Family Responsibilities and Academic Work.” Although more research is needed, at this point we endorse the policies of those institutions that permit a tenure-clock extension of 1 year for childbirth or adoption (maximum of 2 extensions) that is available to both men and women who are the primary caregivers for their children. It is quite possible that some...
women would choose a more traditional career path and not want the tenure clock lengthened.

4. Chairs should explore ways to adjust the balance of the clinical/research ratio such that research productivity can be maintained if women choose to work part-time. For example, if the usual split is 30% clinical/70% research, this ratio should be retained at the 70% level even if a woman works part-time. By contrast, holding the clinical component steady while reducing the research portion effectively elevates the clinical time commitment overall and further handicaps progress in research. We recognize that the maintenance of such a ratio will involve a financial commitment from the institution. Furthermore, it may be difficult for smaller departments to meet clinical care needs and maintain this level of support for research time. In some situations, fathers may also desire, and benefit from, such reduced time.

5. Mentoring, which is important for all junior faculty members in research, is especially critical for women. It is well known that women do not network as well as men and have more difficulty identifying career mentors and finding mentoring opportunities, thus ensuring significant disadvantage for academic advancement. Conversely, studies have shown that women who have a mentor are more likely to be promoted to professor than those without a mentor. There are multiple ways to provide mentoring. In addition to the assigned mentor-mentee relationship, there are successful programs that use collaborative mentoring and facilitated group mentoring that provide a framework for professional development. Regardless of the format, mentoring is a professional activity that is crucial to academic development of all junior faculty members, but particularly women.

Therefore, we encourage all academic institutions to establish policies that provide flexible tenure clocks and modified duties for mothers (or fathers) who are primary caregivers, thus enabling those who are raising families to work part-time and remain on the tenure track. In addition, we strongly recommend that institutions offer formal mentoring programs for junior faculty members, particularly for women entering the research tenure track.

THE DESIRE FOR MORE WOMEN IN LEADERSHIP POSITIONS

The fact that there are so few female chairs in the field of pediatrics is evidence of the fact that women have not risen past the glass ceiling in academia. The percent of female pediatric chairs has not increased appreciably in the past 15 years and has continually hovered at ~10%. Other departments have even fewer female chairs, and there are still very few female deans of medical schools. As the number of women in medicine increases, the paucity of women in senior ranks and leadership positions becomes a liability for the institution. The desire to have more women in leadership positions is not based solely on numbers but, rather, on the desire for diversity in the workplace, opportunities for role modeling, and the ability to have senior-level teams with complementary talents and varying perspectives. To improve these figures, we offer the following suggestions:

1. Institutional search committees should actively look for qualified female candidates for academic leadership positions. In particular, efforts should be made to eliminate unconscious bias against women that undermines the search process to identify and recruit qualified female candidates. Formal bias-awareness training may be helpful in this regard. The National Science Foundation ADVANCE Institutional Transformation Program, the University of Wisconsin, and the Georgia Institute of Technology ADEPT program exemplify the effectiveness of this approach. Furthermore, the American Association of Medical Colleges document on chair responsibilities clearly indicates that chairs must be aware of the subtle challenges that prevent women and minority students and faculty from realizing their full potential and work to address those challenges. Every institution should ensure that there is a climate of acceptance of women in the institution and that all leaders in the institution are held accountable for maintaining such a climate. For example, as part of the chair’s annual performance review, the dean of the medical school could request a status report on the recruitment and retention of female faculty members as well as a plan on how improvements are being implemented. Evaluating chairs on these achievements could be supported by offering chairs educational sessions on improving mentoring in their departments, assistance from the medical school academic affairs office, or access to coaches for themselves or their female faculty.

2. Institutional leadership should support the career development of female faculty members by providing opportunities to attend established leadership-training programs such as the Executive Leadership in Academic Medicine program and the leadership seminars sponsored by the American Association of Medical Colleges. There are 2 different seminars offered: one is for junior faculty who are just starting out in academia, and the other is for women who are at midcareer.

3. Institutions should develop internal or on-campus leadership-training programs. There are many reasons to hold local leadership-training programs. Besides the convenience of not having to travel, institutions can use this venue to have a cadre of participants learn about leadership and develop leadership skills as part of a peer network over time.
4. Institutions should develop internal leadership-mentoring programs and proactively identify qualified women to participate in such programs. Although the desire to have a woman mentor may be expressed by many women, we suggest that having a strong, consistent, and encouraging mentor relationship with a male or female leader is more important than the gender of the mentor in helping women meet their goals. We note that minorities are also underrepresented in leadership positions in academic medicine and would probably benefit from the same attention to mentorship that we are recommending for women.

5. Institutions should include leadership-transition planning as a component of every major academic office, and qualified female candidates should be sought out for these positions when appropriate. Women with families are not always as flexible as men to relocate for the purpose of advancing their careers, which is both a challenge and an opportunity. Universities that identify women with academic potential and groom them for success are more likely to retain them on faculty and make optimal use of their talents.

Therefore, we suggest that institutions hold chairs (and other leaders) accountable for creating a climate of acceptance of women in their departments. Furthermore, although we support the development of on-campus leadership-development programs/academies at all institutions for all faculty, we recommend that such programs specifically target women and other underrepresented minorities. We also advocate active mentoring of talented women to be groomed for leadership positions. Finally, we encourage institutions to sensitize search committees to the presence of unconscious biases that could preclude such women from being chosen for such positions.

CONCLUSIONS
We are entering a new era in the history of medicine when there are equal numbers of men and women becoming physicians. There is a pool of talented women in our midst, and it is our responsibility as leaders to find those individuals and groom them for success. To attract women to academic medicine and ensure that they will have successful careers, we must begin immediately to make thoughtful, meaningful, and even bold changes in academic medical center policies that affect work/home balance. We have addressed 4 issues that we deem to be of the highest importance for the future of women in academics: (1) part-time work options; (2) availability of high-quality day care and child care; (3) flexibility in tenure-track positions; and (4) recruitment of qualified women into leadership positions. We have suggested strategies for addressing these challenges at the departmental and institutional levels but recognize that there may be many other strategies or models that we have not considered. Although the issues raised here apply to all women in academic medicine, the sheer number of women entering pediatrics makes these issues most urgent for our disciplines in the next few years. It is fitting that departments of pediatrics would take the lead in advocating for a new approach to academic careers, one that recognizes the vital role of work/home balance, so that parents who want to be excellent parents as well as outstanding academicians will never have to choose between their children and their careers; both should be possible.

ACKNOWLEDGMENTS
The 2006 AMSPDC women chairs were Barbara Stoll, MD (Emory University); Renee R. Jenkins, MD (Howard University); Rosha McCoy, MD (Medical College of Ohio) (interim); Frances J. Dunston, MD, MPH (Morehouse School of Medicine); Kimberly E. Dow, MD (Queens University); Latha Chandran, MD (State University of New York, Stony Brook) (interim); Ilidy M. Katona, MD, Capt, MC, USN (retired) (Uniformed Services University); Debra H. Fiser, MD (University of Arkansas) (until September 1, 2006); Cheryl Rockman-Greenberg, MD (University of Manitoba); Marianne E. Felice, MD (University of Massachusetts); Valerie P. Castle, MD (University of Michigan); Loretta Cordova de Ortega, MD (University of New Mexico) (interim); Elizabeth R. McAnarney, MD (University of Rochester) (until September 1, 2006); Nina F. Schor, MD, PhD (University of Rochester) (after September 1, 2006); Roberta G. Williams, MD (University of Southern California); Ellen R. Wald, MD (University of Wisconsin); Bonita Stanton, MD (Wayne State University); and Margaret K. Hostetter, MD (Yale University).

The women chairs of the AMSPDC appreciate the support of our male colleagues who encouraged us to write and publicize the concerns presented in this article. We particularly recognize the contributions and encouragement of F. Bruder Stapleton, MD, President of AMSPDC, 2005–2007, and we extend our gratitude to Richard Behrman, MD, who chaired the FOPO Committee and spoke frequently and passionately about the need for change. We also thank Jean Bartholomew for her helpfulness and assistance to all of us in the AMSPDC.

REFERENCES
STENTING OPENS JAMMED ARTERIES IN THE BRAIN

“By pushing a tiny mesh cylinder called a stent through blood vessels leading from the groin to the head, doctors can prop open narrowed arteries in the brain much as they do in the heart, several new studies show. A brain artery that’s partially blocked because of arteriosclerosis is a stroke waiting to happen. . . . Scientists have adapted stents to fit brain arteries, which are smaller and more fragile than the arteries serving the heart. Two studies presented last week at the 2007 International Stroke Conference in San Francisco, along with a trial reported last year, indicate that the still-experimental brain stents might work as well or better than drugs and have fewer adverse effects.”

Science News. 2007;171:99

Noted by JFL, MD
COMMENTARY

Circumcision in the Time of HIV: When Is There Enough Evidence to Revise the American Academy of Pediatrics’ Policy on Circumcision?

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.

There have been 3 recent studies in Africa, involving >10 000 men, that have demonstrated a marked protective effect of male circumcision with respect to the acquisition of HIV infection.1–3 The protective effect was 60% in each of the 3 trials. Furthermore, men who were circumcised were no more likely than uncircumcised men to engage in risky sexual behavior.2–4 Commentaries5,6 appearing in the same issue that published the 2 latest trials strongly affirm circumcision as a means of preventing HIV infection. Reviews of the literature7,8 have concluded that there is substantial evidence to support the conclusion that circumcision significantly reduces the rate of HIV infection, and one review concluded that “male circumcision is the most compelling evidence-based preventative strategy to emerge since the results of mother-to-child transmission clinical trials.”9

In the United States in 2005 there were 1434 new cases of HIV infection in children and young adults 19 years of age or less10 and 453 new cases of syphilis in the same age group,11 whereas the prevalence of human papillomavirus infection among females 14 to 19 years of age who were surveyed in 2003–2004 was 24.5%.12 Although HIV infection occurs much less frequently in the United States when compared with the developing world, it still represents a substantial problem.

Circumcision also protects against certain other sexually transmitted diseases (STDs). Authors of the first systematic review and meta-analysis of the association of male circumcision with ulcerative STDs (syphilis, cancrum, and genital herpes)13 concluded that circumcised men are at lower risk of acquiring cancrum and syphilis than uncircumcised men. There is also compelling evidence that male circumcision protects against human papillomavirus infection and, hence, cervical14–19 and penile cancer.20

The American Academy of Pediatrics (AAP) issued its most recent policy on newborn circumcision in 199921 and reaffirmed its conclusion in 200022 and 2005.23 The most recent statement concludes that although there are “potential medical benefits... these data are not sufficient to recommend routine neonatal circumcision.”22 As discussed in 2 commentaries critical of the AAP’s policy,24,25 the evidence for the beneficial effects of circumcision seem to have been underappreciated by the authors of the policy statement. The benefits include virtual elimination of penile cancer, as well as a marked decrease in balanoposthitis, phimosis, paraphimosis, and penile dermatosis.25 It has also been pointed out that the AAP listed 6 evidence-based benefits and only one minor risk (a surgical complication rate of 0.2%–0.6%).25

There is little argument that circumcision reduces the incidence of urinary tract infection (UTI) in infants26; the only question involves the magnitude of its beneficial effect. Some suggest that this benefit only applies to boys at high risk of UTI,27 whereas others point out that the cost/benefit ratio of preventing renal scarring, which may occur in 18% of boys who present with UTI, may make the procedure cost-effective.28

In 2004, our colleagues in obstetrics and gynecology stated that “a consensus is forming that circumcision offers protection against UTI, penile cancer, cervical cancer, genital ulcer disease, and HIV.”26 The authors of this article, as well as others,21 discussed the various ways in which pain control during neonatal circumcision can be achieved and also concluded, as have others,25 that there

Abbreviations: STD, sexually transmitted disease; AAP, American Academy of Pediatrics; UTI, urinary tract infection

Opinions expressed in these commentaries are those of the authors and not necessarily those of the American Academy of Pediatrics or its Committees.

www.pediatrics.org/cgi/doi/10.1542/peds.2007-0739

doi:10.1542/peds.2007-0739

Accepted for publication Mar 13, 2007

Address correspondence to Joseph D. Dickerman, MD, Department of Pediatrics, University of Vermont College of Medicine, D301 Given Building, Burlington, VT 05405. E-mail: joseph.dickerman@uvm.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics.
is no increase in sexual dysfunction after circumcision. They further suggested that even after parents are given the most current information on the risks and benefits of circumcision, their decision is often based on social, cultural, religious, and racial factors, as well as the circumcision status of the father. It was also opined that “some of the medical literature about the procedure suffers from authors who put the fury of debate above the science.” It is clear that, rather than using evidence-based data, some in the medical community who oppose neonatal circumcision use similar factors on which to base their opinions. Parents should always have the right to choose whether to have their neonate circumcised. However, they must be presented with accurate, unbiased, evidence-based data. A revised AAP policy that reflects the recent findings described above would provide health care professionals and parents with an appropriate tool to allow them to arrive at an informed decision.

It is very disturbing to note that the prevalence of circumcision has declined in the United States from 91% in the 1970s to 83% in the 1980s. From 1999–2000, it was 79%. In this age without an AIDS vaccine when many individuals, especially teenagers, practice risky sexual behavior and a significant number of people do not use condoms because of religious beliefs, lack of appropriate education, inability to afford them, or difficulty in acquiring them, circumcision may offer the best method for protection against certain STDs, especially HIV.

I firmly believe that there is enough new information to prompt a revised AAP policy statement regarding neonatal circumcision, considering the very significant beneficial effects and the very minor risks associated with the procedure.

REFERENCES

5. Newer approaches to HIV prevention. Lancet. 2007;369:615
General Recommendations on Immunization: Recommendations From the Advisory Committee on Immunization Practices


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.
Inhalant Abuse
Janet F. Williams, MD, Michael Storck, MD, and the Committee on Substance Abuse and Committee on Native American Child Health

ABSTRACT
Inhalant abuse is the intentional inhalation of a volatile substance for the purpose of achieving an altered mental state. As an important, yet-underrecognized form of substance abuse, inhalant abuse crosses all demographic, ethnic, and socioeconomic boundaries, causing significant morbidity and mortality in school-aged and older children. This clinical report reviews key aspects of inhalant abuse, emphasizes the need for greater awareness, and offers advice regarding the pediatrician’s role in the prevention and management of this substance abuse problem.

TYPES OF CHEMICALS AND PRODUCTS ABUSED
The term “inhalant” encompasses a wide range of pharmacologically diverse substances that readily vaporize. Most other substances of abuse are classified by grouping together substances that share a specific central nervous system action or perceived psychoactive effect, but inhalant substances that are abused are grouped by having a common route of drug use. Inhalant abuse, sometimes referred to as solvent or volatile substance abuse, can be better understood when the expansive list of inhalants is classified into 3 groups on the basis of what is currently known pharmacologically: group I includes volatile solvents, fuels, and anesthetics; group II includes nitrous oxide; and group III includes volatile alkyl nitrites (Table 1).

Inhalant abusers use volatile products that are capable of producing a quick and generally pleasurable sensory experience, or “high,” with rapid dissipation and minimal “hangover” symptoms. Inhaled substances are widely available, convenient, inexpensive, easily concealed, and legal for specific intended uses but are intentionally misused by abusers. Many of these qualities are important factors that promote use in a young age group, because children have less sophisticated resources for acquiring alternative substances of abuse. The most commonly abused inhalants are the group I aliphatic, aromatic, or halogenated hydrocarbons found in thousands of commonly used and readily available consumer products. Virtually any hydrocarbon can have mind-altering effects when inhaled in large doses. Nitrous oxide or “laughing gas” is diverted from medical or dental anesthesia use and sold in balloons for inhalation or is simply inhaled from whipped cream aerosol cans. Alkyl nitrites or “poppers” are also abused; prototypically, amyl nitrite ampules intended to treat angina are “popped” open and inhaled.
TABLE 1 Pharmacologic Classification of Inhalants: Selected Common Street Names and Chemical Content of Product Examples

| I. Volatile solvents, fuels, and anesthetics (air blast, discorama, hippie crack, medusa, moon gas, oz, poor man’s pot) |
| Solvents: toluene, acetone, methylene chloride, ethyl acetate, TCE in paint thinner, paint and polish removers, correction fluid, and felt-tip marker fluid; TCE and tetrachloroethylene in degreasers, spot removers, and dry-cleaning fluids; toluene, hexane, TCE, ethyl acetate, and methyl chloride in glues and rubber cement; propellants and solvents such as butane, propane, chlorofluorocarbons, hydrocarbons in aerosol spray paint, computer/electronics-cleaning spray, spray deodorant, hair spray, vegetable-oil cooking spray, air-fresher spray, fabric-guard spray, and anesthetic sprays |
| Fuels: butane or propane lighters or pressurized fuel tanks, gasoline, racing car octane boosters, refrigerants |
| Anesthetics: ether, halothane, enflurane, ethyl chloride |

II. Nitrous oxide (laughing gas, buzz bomb, shoot the breeze): diverted medical anesthetic, whipped-cream dispenser charger (whippets), whipping-cream aerosol

III. Volatile alkyl nitrites (poppers, snappers, boppers, pearls, amys [isooamy nitrite diverted from medical use], quicksilver, and brand/slang terms Rush, Bolt, Thrust, Climax, Locker Room): videocassette-recorder head cleaner, liquid aroma/liquid incense air fresheners or room odorizers (mostly cyclohexyl nitrite), isobutyl nitrite or butyl nitrite, isopropyl nitrite

TCE indicates 1,1,1-trichloroethane.

EPIDEMIOLOGY

Inhalant abuse occurs throughout the world, in industrialized nations as well as developing countries. Several studies have helped define the epidemiology of volatile substance abuse in the United States.3–11 The peak age of inhalant abuse is 14 to 15 years, with onset in children as young as 5 or 6 years of age. Use typically declines by 17 to 19 years of age but can continue into adulthood. Use by adults may predominate under particular circumstances, such as when certain occupations make abusive solvents, propellants, or anesthetics readily available. Inhaled nitrites have a long history of being abused in certain social settings, particularly when men have sex with men.1 The type, frequency, and method of volatile substance abuse vary widely in relation to the age of the abuser, geographic region, and ease of availability.

The National Survey on Drug Use and Health (NSDUH), an annual survey of drug use in the general US noninstitutionalized civilian population 12 years and older, has documented inhalant abuse initiation by both adolescents and adults.10 In 2005 and similar to previous years, 72.3% of the 877 000 new volatile substance abusers aged 12 to 49 years were younger than 18 years, with a mean age of 16.1 years. Since 2002, no significant change has occurred in the number of inhalant initiates, the average age of first use, or the rates of inhalant abuse by either youth or adults. This survey again showed no significant male-female difference in lifetime prevalence of inhalant abuse in the 12- to 17-year age group but confirmed a greater prevalence of inhalant abuse by men in the 18- to 25-year age group, suggesting that sustained use of inhalants is more common in males. Comparison of 2004 NSDUH findings with the Canadian Addiction Survey, a telephone survey in early 2004 of Canadian household residents 15 years and older, showed that US residents were 7 times more likely (9.5%) than Canadians (1.3%) to have ever used inhalants.12

According to the National Institute on Drug Abuse and the University of Michigan’s annual Monitoring the Future (MTF) survey results11:

- Prevalence of lifetime inhalant use (“ever used”) among 12th-graders has ranged between 10.3% in 1976 (when first included in the survey) and 18.0%, the 1990 peak. The 2006 rate of 11.1% has been stable since 2002.
- Since 1979, prevalence has additionally been reported after adjustment for the underreporting of amyl and butyl nitrite use. Adjusted lifetime prevalence figures remained at or above 17.0% until 1997 before steadily declining and stabilizing near the 2004 low of 11.4%. Adjusted lifetime and annual-use (“at least once in the past year”) rates for 12th-graders (11.5% and 4.7%, respectively, in 2006) are among the lowest levels in survey history.
- Roughly similar declines in prevalence of inhalant use have been documented in the 8th- and 10th-grade age groups, which the MTF survey has included since 1991 but does not adjust for possible nitrite use.
- Recent data on perceived harmfulness may be portentous. Since 2001, the percentages of 8th- and 10th-graders who indicated that they “think people risk harming themselves (physically or in other ways) if they try inhalants once or twice” or “try inhalants regularly” have decreased. Past research has shown that decreasing perceived risk of using a drug often precedes an upswing in use of that drug.

MTF survey results have consistently shown that the reported prevalence of inhalant use by 8th-graders has been, on average, approximately 2% to 3% greater than that of 10th-graders, which runs approximately 2% greater than that of 12th-graders. This pattern, which is opposite that of nearly all other abused substances, may simply reflect that early experimentation with this easily acquired drug class is greater in younger age groups, that older students may fail to report inhalant use that occurred in earlier grades, or that many 8th-grade inhalant abusers subsequently drop out of school and are, therefore, no longer included in the survey population.6,13,14 Research has shown that inhalant use often occurs in conjunction with other risk behaviors and that higher rates of inhalant abuse occur among children who have poor grades or have dropped out of school compared with classmates who remain in good standing at school.6,13–15

Although inhalant abuse is more prevalent among
physiologically isolated and socioeconomically disadvantaged populations, it crosses all demographic boundaries and occurs in rural as well as urban settings and among all ethnic groups in the United States.\textsuperscript{4,6–11} Important universal factors that promote initial experimentation with inhalants and their continued use include peer use and low perceived harm from use.\textsuperscript{8,11,16} Inhalant use is often associated with impoverished living conditions, delinquency, criminal behavior, incarceration, depression, suicidal behavior, greater antisocial attitudes, family disorganization and conflict, or a history of abuse, violence, or other substance abuse, including injection drug use.\textsuperscript{4,13,15–21}

MTF surveys have documented lower rates of past-year inhalant use among Hispanic compared with white individuals, with the lowest rates consistently among black individuals.\textsuperscript{11} Similarly, most other studies have found rates of inhalant use by Hispanic youth to be the same as or lower than use by non-Hispanic white youth.\textsuperscript{4,8,10,11} NSDUH data consistently show that rates of inhalant abuse by Asian American youth are among the lowest.\textsuperscript{4,10} Inhalant use has been seen as a particularly serious problem among American Indian/Alaska Native youth for many years.\textsuperscript{6,10,22} Research to discern factors contributing to inhalant abuse suggests that adverse socioeconomic conditions, including isolation and lower educational levels rather than racial or cultural factors per se, account for the reported higher rates among these minority populations. Eskimo children 10 to 19 years of age living in 14 isolated Bering Strait villages reported a lifetime prevalence of inhalant use of 48%.\textsuperscript{23} American Indian youth living on reservations have been shown to have higher rates of inhalant abuse than do either American Indian youth living off reservations or non-Hispanic white youth.\textsuperscript{22} Paralleling the decreasing inhalant abuse shown by the MTF studies until 2004, a promising downward trend has also been demonstrated since 1995 by the annual survey of American Indian students living on reservations.\textsuperscript{8} Because use of other drugs during much of this time period also decreased, substitution of other drug use for solvent use was not felt to explain the trend. Despite these epidemiologic data, inhalants remain among the least studied groups of abused substances. Much research is needed to understand all aspects of inhalant use, including the neuropharmacologic effects and psychosocial correlates.

**MECHANISM OF ABUSE AND IMMEDIATE EFFECTS OF INHALANTS**

Inhalants are abused through a variety of methods, and many “street” terms for this activity have been generated, such as (glue) sniffing, snorting, huffing, glading, and dusting. Product fumes are usually inhaled through the mouth (huffing) or nose (sniffing or snorting) from the original container. Abusers may also inhale vapors from a chemical-saturated rag held to the face or stuffed in the mouth, which is also called huffing. Some aerosols are sprayed directly into the mouth or nose, and volatile solvents can be applied onto the nasal mucosa or a nearby surface such as fingernails or a shirt collar or cuff and then inhaled. “Glading” refers to the inhalation of air-freshener aerosols, whereas a recently coined term, “dusting,” refers to the abuse of aerosol computer and personal electronics cleaning products by placing the canister straw into the mouth or nose. Familiar and innocuous containers are often used to help conceal inhalant abuse (eg, inhaling spray paint fumes out of a soft drink can or nitrous oxide–filled balloons). A paper or plastic bag containing the inhalant can be held to the mouth and nose or over the head (“bagg ing”).

Unusual fads in inhalant abuse products or methods have been reported, such as heating volatile substances and inhaling the released vapors, as has been done with certain fertilizers or “snotballs” of rubber cement.\textsuperscript{24} Mothballs have been abused by bagging or chewing.\textsuperscript{25} Products combining inhalants with nonrespirable toxic ingredients, such as antiperspirants containing the toxic compound aluminum chlorohydrate, can be bubbled through water.\textsuperscript{26} Combined alcohol and inhalant abuse by drinking “ocean” has been a periodic problem on and near some southwest American Indian reservations. Mixing water or mouthwash with the contents of a hairspray can, usually AquaNet, containing ethanol, methanol, and a propellant, produces foamy “ocean water” and combined toxicities.\textsuperscript{27}

Inhalants are readily absorbed through the lungs, with immediate and brief effects, and then relatively rapidly metabolized predominately through the cytochrome P450 system of the liver. Inhalants, except nitrates, are depressants that act directly on the central nervous system through a wide range of mechanisms yet to be completely elucidated.\textsuperscript{1,28} As a group, inhalants most resemble alcohol, whereby different cellular mechanisms are responsible for myriad pharmacologic and toxicologic effects. Opiate receptor involvement likely plays a role in the analgesic effects of nitrous oxide, but there is evidence for γ-aminobutyric acid (GABA)–mediated behavioral effects.\textsuperscript{1} Volatile hydrocarbons also have GABAergic effects and a possible role in the inhibition of glutamatergic neurotransmission involving N-methyl-D-aspartate (NMDA) receptors.

The immediate effects of inhaling volatile solvents, fuels, anesthetics, or nitrous oxide are similar to the early stages of anesthesia. The user feels an initial stimulating “rush,” then is light-headed, disinhibited, excitable, and prone to impulsive behavior. Intoxication lasts only a few minutes but can be extended for several hours by breathing inhalants repeatedly. Slurred speech, dizziness, diplopia, ataxic gait, and disorientation occur as the inhalant dose increases. Euphoria is followed by drowsiness, a lingering headache, and sleep, particularly after repeated cycles of inhalation. Visual hallucinations

---

**References**


are possible with prolonged use. Coma is unusual, because as the user becomes drowsy, exposure to the inhalant is usually terminated before a large enough dose is absorbed to cause severe neurologic and respiratory depression. Mucous membrane irritation may manifest as rhinorrhea, epistaxis, sneezing, coughing, excess salivation, and conjunctival injection. Some patients experience nausea, vomiting, diarrhea, abdominal cramps, dyspnea, or wheezing.3,29

Nitrites significantly differ pharmacologically from other inhalants, because instead of direct central nervous system effects, they primarily cause vasodilation and smooth muscle relaxation. The sensations of floating and increased skin tactility as well as warmth and throbbing occur within 10 seconds of inhalation but diminish within 5 minutes. Nitrite abuse may result in tachycardia, flushing, blurred vision, headache, lightheadedness, significant hypotension, syncope, and sufficient methemoglobinemia to cause cyanosis and lethargy.1,3,29 Other inhalants are used to alter mood, but nitrites are inhaled to enhance sexual feelings, penile engorgement, and anal sphincter relaxation to intensify sexual experience.3,29

MORBIDITY AND MORTALITY

Patterns of inhalant abuse are similar to those of other substance abuse, and users can generally be described as experimenters, intermittent users, or chronic inhalant abusers. Similarly, morbidity and mortality increase as frequency of use increases, with the important exception that “sudden sniffing death syndrome” is a risk during any use, even during initial experimentation. In 1 study, 22% of inhalant abusers whose deaths were attributed to sudden sniffing death syndrome had no history of previous inhalant abuse.30 Sudden sniffing death syndrome is the leading cause of fatality related to inhalant abuse.

Bass31 originally described sudden sniffing death and elucidated its pathophysiology. Hydrocarbons and other inhalants “sensitize” the myocardium to epinephrine, and when this hormone is produced in response to any of a number of stimuli, most commonly sudden stress or fright, a fatal cardiac arrhythmia can result. Instead of truly sensitizing the cells, volatile substances stabilize myocardial cell membranes to depolarization. Because of variable individual myocardial cell response and the complex way that myocardial electrical impulses are propagated, greater cell stability actually blocks electrical impulse conduction and increases arrhythmia risk. During inhalant use, arrhythmias can occur even with normal epinephrine concentrations, but an adrenaline surge, such as when hallucinating or when discovered by or running from an authority figure, increases the risk.32 Sudden sniffing death can occur during inhalation or in the subsequent few hours, because a volatile substance dissolved in lipid-rich cell membranes dissipates relatively slowly.32 This unpredictable and unpreventable type of death leaves no specific macroscopic or microscopic postmortem features, so no cause can be identified at autopsy.

Death caused by inhalant abuse can also occur through a variety of other mechanisms but is usually attributable to an acute and related event, most likely suffocation, aspiration, or accidental injury (Table 2). From 1981 to 1985 in Britain, suffocation, aspiration, and accidental injury each accounted for approximately 15% of deaths attributable to inhalant abuse, and the remaining 56% of deaths were attributed to sudden sniffing death syndrome.30 Suffocation occurs when the mode of use involves inhalation through the nose and mouth from a plastic bag, which may occlude the airway if the user loses consciousness. The risk of death caused by aspiration, usually of vomitus, is similar to that for alcohol or other depressants and is related to the combination of a decreased level of consciousness and the loss of protective airway reflexes. While under the influence of inhalants, users become less inhibited as well as less alert and oriented, which can promote their engaging in risk behaviors and lead to accidental injury and death, such as from a motor vehicle crash, drowning, fire, a jump or fall from heights, or hypothermia from exposure to the elements.

The United Kingdom, with a population approximately one fifth that of the United States, has been the only major part of the Western world to track in a systematic way deaths associated with volatile substance abuse. Since 1999 legislation banned the sale of butane cigarette lighter refills to youth younger than 18 years, there has been a significant drop in inhalant use deaths in both this age group and older individuals. The 2003 volatile substance abuse–related death total of 51 was the lowest number recorded for the United Kingdom since 1983. Of the 9 individuals younger than 18 years who died, 6 did so in relation to inhalation of butane lighter refills, compared with 15 of the 24 deaths in this age group in 2002.33

Three reports shed light on the US inhalant abuse mortality rate. The Toxic Exposure Surveillance System (TESS) database of the American Association of Poison Control Systems showed 63 deaths in 11 670 cases of

---

**TABLE 2 Causes of Death From Inhalant Abuse**

<table>
<thead>
<tr>
<th>Type</th>
<th>Causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute</td>
<td>Direct causes: immediate or “postponed” sudden sniffing death syndrome; methemoglobinemia</td>
</tr>
<tr>
<td></td>
<td>Indirect causes: suffocation, aspiration, trauma, drowning, fire, other</td>
</tr>
<tr>
<td>Delayed</td>
<td>Cardiomyopathy</td>
</tr>
<tr>
<td></td>
<td>Central nervous system toxicity: toluene dementia and brainstem dysfunction</td>
</tr>
<tr>
<td></td>
<td>Hematologic: aplastic anemia, leukemia</td>
</tr>
<tr>
<td></td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td></td>
<td>Renal toxicity: nephritis, nephrosis, tubular necrosis</td>
</tr>
</tbody>
</table>

---
intentional inhalant abuse reported from 1996 to 2001 to poison-control centers nationwide. Actual mortality rates are likely greater, as evidenced by extrapolation from 2 studies that examined state death records that mentioned inhalants as a contributing cause of death at any age. These studies found 39 deaths in Virginia from 1987 to 1996 and 144 deaths in Texas from 1988 to 1998. In Virginia, 70% of those who died were 22 years and younger, and in Texas, 28.7% of victims were 8 to 17 years of age. Of the inhalant abuse cases reported to the TESS, 54% were in youth 13 to 19 years of age, 15% were in children 6 to 12 years of age, and 0.4% were in children 5 years and younger. The 63 fatalities occurred almost exclusively in adolescents and young adults. Three types of inhalants were associated with the majority of deaths reported to the TESS: gasoline (45%), air fresheners (26%), and propane/butane (11%). These same group I inhalants (Table 1) were associated with the majority of deaths in both Virginia and Texas, particularly fuels including refrigerants and various solvents.

There is as diverse a list of possible sequelae of chronic inhalant abuse as there is diversity in the types of volatile solvents, fuels, and anesthetics used and the dose and frequency of exposure. If chronic solvent abuse is terminated, there is remarkable reversibility of many of the pathologic effects, but compared with other organ systems, the nervous system has less regenerative capacity. Of all biological membranes, myelin has the highest fat content at 75%, and neuronal membranes may contain up to 45% lipid. The primary consequence of frequent and longer-term inhalant use over months to years is chronic nervous system absorption of these highly lipophilic substances and significant nervous system damage, resulting in muscle weakness, tremor, peripheral neuropathy, cerebellar dysfunction, chronic encephalopathy, and dementia, including mood changes.37–39 Computed tomography has demonstrated a loss of brain mass, and magnetic resonance imaging has shown white-matter degeneration and subcortical abnormalities, particularly in the thalamus, basal ganglia, pons, and cerebellum.40 Cognitive impairment has been reported with deficits found in memory, attention, auditory discrimination, problem-solving abilities, visual learning, and visual-motor function.29,38 A limitation of the few studies that have investigated cognitive and neuropsychiatric functioning of inhalant abusers is that most of them have not adequately demonstrated that the impairments were not premorbid deficits. Most of the acute neurologic, neuropsychiatric, and cognitive sequelae of volatile solvent abuse seem to be reversible, but the resolution of chronic symptoms is much slower and less complete.29,38

Other causes of morbidity and mortality are related to the specific volatile chemical(s) used, associated health risk behaviors, drug-drug interactions, or additional material(s) found in the various inhaled products. Toxic effects attributed to specific chemicals include an ichthyosis-like dermatitis on the extremities,25 decreased visual acuity,41,42 sensorineural hearing loss,42 cardiomyopathy,43 toxic hepatitis,44 distal renal tubular acidosis,45 metabolic acidosis,46 leukemia,47 and aplastic anemia.48 There is evidence that tolerance, dependence, and withdrawal symptoms can occur, and reported morbidities also include toluene embryopathy and neonatal withdrawal.49–52 Lung damage from paint pigments, lead poisoning from leaded gasoline, and other such toxicities have been reported when an inhalant contains another potential toxin.53 Inhalant abuse is associated with the abuse of other substances, including pharmaceuticals, alcohol, tobacco, and illicit drugs, which can obscure the diagnosis of inhalant abuse and increase potential morbidity.21,54–55 Combining other drugs with inhalants expands the potential for risk behaviors, altered drug metabolism, and drug-drug interactions, including potentiation of drug effects, particularly depressant effects. High flammability and accidental combustion of volatile agents have led to burns and other fire-related injuries.3,24

Chronic nitrous oxide abuse causes short-term memory loss and peripheral neuropathy, which subside with discontinuation of the abuse.29 Peripheral neuropathy results from nitrous oxide inactivating vitamin B₁₂ and mediating a pernicious anemia-type syndrome, which includes anemia, leukopenia, sensorimotor neuropathy, and posterior/lateral column spinal cord disease.29 Nitrite inhalation has been associated with hemolytic anemia in patients with glucose-6-phosphate dehydrogenase deficiency.56 Because nitrites are abused mainly for their sensory and sexual effects, use may promote higher-risk sexual practices, facilitate transmission of sexually transmitted infections, and result in drug interactions, such as with sildenafil.3 Chronic abuse of volatile alkyl nitrites has documented hematologic and immune system effects without associated cognitive deficits.1,29,56

**TABLE 3** Major Neurotoxic Consequences of Inhalant Abuse

<table>
<thead>
<tr>
<th>Cerebellar ataxia</th>
<th>Cranial neuropathy: usually cranial nerves V and VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encephalopathy: acute or chronic</td>
<td>Multifocal: both cortical and subcortical central nervous system damage, both central nervous system and peripheral nerve effects</td>
</tr>
<tr>
<td>Optic neuropathy: visual loss</td>
<td>Parkinsonism</td>
</tr>
<tr>
<td>Peripheral neuropathy</td>
<td></td>
</tr>
</tbody>
</table>

**DETECTION OF INHALANT ABUSE**

Inhalant abuse may not readily come to the attention of others, including pediatricians, because signs and symptoms of use are often subtle. Abuse of inhalants should
be suspected when a cache of a potential inhalant is discovered or when products with abuse potential are found stored in unusual locations, such as cans of gasoline or spray paint under a youth’s bed. Changes in an adolescent’s behavior, including apathy, malaise, poor appetite, a significant shift in choice of friends or activities, or an unexplained drop in school grades, can also be signs of inhalant abuse. Those who are chronic and heavy inhalant abusers may be identifiable because of their combination of poor hygiene and grooming, weight loss from decreased caloric intake, and chronic complaints of, for example, fatigue, rhinitis, conjunctivitis, recurrent epistaxis, and oral or nasal ulcerations. Chronic neuropsychiatric changes, such as confusion, poor concentration, depression, irritability, hostility, or paranoia, may predominate. Symptoms of other organ system toxicities from long-standing inhalant use may also bring the abuser to medical attention.

Inhalant abusers may present with obvious intoxication and evidence of use, such as a conspicuous odor of the inhalant. This chemical odor is most often present because the abuser excretes a significant proportion of the absorbed dose when exhaling, and the odor can persist on the breath for many hours. If the abused product spilled onto clothing during use or was intentionally put on clothing, the odor may also persist, and clothing stains or paint may be found. Paint or glitter may also be seen on the abuser’s face or hands, or there may be a “huffer’s rash,” classically a perioral or perinasal dermatitis with pyoderma. Contact with inhalants dries the skin and leads to small cracks, which allow bacteria to enter. The dermatitis may look like a non-specific contact hypersensitivity reaction or perioral eczema or, with nitrite use, may have a yellow crust at specific contact hypersensitivity reaction or perioral eczema. When considered in certain clinical contexts, abnormal nontoxicologic laboratory results, such as elevated liver enzymes, may arouse or confirm a suspicion of inhalant abuse. Blood and other tissues, usually brain or liver, can be tested by specific gas chromatography technique when inhalant detection is necessary, such as in a fatality. Specific urine drug testing is sometimes useful as part of the treatment-compliance plan when benzene, toluene, or a similar agent has been chronically abused, because major urinary metabolites (phenol and hippuric acid, respectively) are detectable when there has been a high level of use.

**INHALANT ABUSE PREVENTION AND MANAGEMENT CONSIDERATIONS**

As with other types of substance abuse, the most effective way to curtail use is through broad prevention efforts, particularly primary prevention through education paired with skills-building. Developmentally and culturally sensitive educational strategies should be implemented, such as those implemented in many American Indian communities through a prevention initiative in conjunction with the American Indian Institute at the University of Oklahoma.

Limiting the availability of volatile substances is impractical, because they constitute products that are universally available and legal and have legitimate uses. Restricting the availability of some of these products, such as the United Kingdom ban on the sale of butane lighter refills to youth, can be successful but may also promote the use of other more-available products or create a black market for the restricted products. Adding a noxious chemical to the product to prevent misuse was tried with plastic glue and found to be ineffective, because multiple products would require such adulterants, abusers switch products, and legitimate consumers and product efficacy might be adversely affected. Reformulating the product by replacing the hydrocarbon with other chemicals has occurred when economically feasible and when product efficacy could be maintained. Product warning labels can alert the public to inhalant dangers but may also promote easy identification of abusable substances. Most states have laws making the use of inhalants or sale to minors illegal, and although difficult to enforce and of yet-unproven efficacy, such laws serve as a reminder that society condemns inhalant abuse.

Most acutely intoxicated inhalant abusers do not seek medical attention, and only when intoxication is life-threatening or has led to serious injury will an abuser present to the emergency department. Acute medical management of inhalant abuse starts with applying the “ABCs” of life support to assess and stabilize the patient and address any specific acute injury or toxicity, such as combating methemoglobinemia by administering intravenous methylene blue. Hydration and cardiorespiratory status should be monitored closely. Myocardial sensitization by inhalants necessitates a calm and supportive environment in which the use of pressor medications and bronchodilators are relatively contraindicated. No medications reverse acute inhalant intoxication or have
been found to be helpful with dependence or withdrawal symptoms. Decontamination of the patient’s clothing and skin may be indicated. Laboratory testing can help monitor oxygenation and hematologic status and detect other substances being abused. Testing for organ-system damage should be considered only when there is a history of regular and long-term inhalant use. After acute stabilization, comprehensive medical care includes documenting a detailed history and physical examination and specifically evaluating the patient’s mental health, substance abuse history, and psychiatric needs so that appropriate inpatient or outpatient interventions can be initiated.\textsuperscript{28,29,57}

Little research exists concerning treatment needs and successful treatment modalities specific to inhalant users, so clinicians rely on applying methods that are used to treat other addictive disorders, such as cognitive-behavioral therapy, multisystem and family therapy, 12-step facilitation, and motivational enhancement techniques.\textsuperscript{62} Inhalant abusers seem to respond best to a treatment program that includes an extended detoxification or “treatment readiness” period of 4 or more weeks, during which basic supportive care and general orientation are emphasized. If sufficient time is not allowed, individuals seem incapable of engaging in the treatment program.\textsuperscript{63} Some treatment facilities have used a peer-advocate system for patients, which seems to offer a nonthreatening and supportive treatment approach.\textsuperscript{64} Neuroleptics and other forms of pharmacotherapy are usually not useful in the treatment of inhalant abusers except to address comorbid conditions. Increasing personal and ethnic self-identity through role-modeling has been suggested as helpful in treating some groups of inhalant abusers, and positive cultural identification has been shown to be important in American Indian/Alaska Native populations.\textsuperscript{65} Treatment challenges are posed by the diversity of abused inhalants and user populations, comorbid psychopathology, psychosocial problems, polydrug use, and the physiologic and neurologic effects of inhalant abuse.\textsuperscript{65,66} Treatment of longer-term inhalant users is hindered by the fact that there are few programs designed specifically for inhalant abuse treatment, access to care may be limited, providers generally have a pessimistic view about users’ neurologic damage and chance for recovery, and providers often lack sufficient knowledge and training about inhalant abuse, inhalant users, and their treatment needs.\textsuperscript{64} Although the principles of effective substance abuse treatment in general apply to inhalant abuse treatment, any treatment regimen must address the many clinical, emotional, social, academic, pharmacologic, neurocognitive, cultural, and demographic factors that make this type of substance abuse unique. Treatment strategies are still under development, and additional research is needed to identify effective strategies for the treatment of children and adolescents who use inhalants.

CONCLUSIONS AND ADVICE
The American Academy of Pediatrics has established recommendations\textsuperscript{66} regarding the pediatrician’s role in the prevention, identification, and management of substance abuse and advises the following to promote that role with regard to inhalant use by youth.

1. Pediatricians are encouraged to:
   - be aware that inhalant abuse occurs in all patient populations, including their own;
   - be knowledgeable about the epidemiology of inhalant abuse, particularly about local and regional trends, as well as resources, such as the telephone number 1-800-222-1222 to contact the nearest poison-control center;
   - be knowledgeable about health consequences of inhalant abuse and, in particular, about unique clinical features such as central nervous system damage and sudden sniffing death syndrome;
   - assist in educating children, adolescents, parents, teachers, media representatives, and vendors of volatile substances regarding inhalant abuse prevention and the health risks of inhalant use; and
   - serve as a community resource regarding inhalant use awareness, prevention, detection, and management using national and local community resources such as the National Inhalant Prevention Coalition (1-800-269-4237 or www.inhalants.com), an information and referral clearinghouse.

2. Inhalant abuse education can be included in all substance abuse prevention curricula in the primary and secondary grades, using approaches that effectively warn against the dangers of inhalant use yet do not inadvertently introduce youth to available substances with abuse potential.

3. Widespread accessibility and use of research-based resources such as National Institute on Drug Abuse publications (available at www.drugabuse.gov/Drug-Pages/Inhalants.html) are encouraged.

4. Increased research efforts to evaluate prevention and treatment approaches specific to inhalant abuse and to identify those with efficacy are needed.

COMMITTEE ON SUBSTANCE ABUSE, 2005–2006
Alain Joffe, MD, MPH, Chairperson
Marylou Behnke, MD
John R. Knight, MD
Patricia K. Kokotailo, MD, MPH
Tammy Sims, MD, MS
Janet F. Williams, MD\textsuperscript{*}
REFERENCES

17. Howard MO, Jenson JM. Inhalant use among antisocial youth: prevalence and correlates. Addict Behav. 1999;24:59–74
22. Beauvais F. Comparison of drug use rates for reservation In-


36. Lolin Y. Chronic neurological toxicity associated with exposure to volatile substances. *Hum Toxicol.* 1989;8:293–300


44. Fregert S, Poulsen J, Trulsson L. Yellow stained skin from sodium nitrite in an itching agent. *Contact Dermatitis.* 1980;6:296


Maltreatment of Children With Disabilities

Roberta A. Hibbard, MD, Larry W. Desch, MD, and the Committee on Child Abuse and Neglect and Council on Children With Disabilities

ABSTRACT

Widespread efforts are being made to increase awareness and provide education to pediatricians regarding risk factors of child abuse and neglect. The purpose of this clinical report is to ensure that children with disabilities are recognized as a population that is also at risk of maltreatment. Some conditions related to a disability can be confused with maltreatment. The need for early recognition and intervention of child abuse and neglect in this population, as well as the ways that a medical home can facilitate the prevention and early detection of child maltreatment, are the subject of this report.

INTRODUCTION

The maltreatment of children, including those with disabilities, is a critical public health issue. For purposes of this report, the terms “disability” and “special health care needs” include the full spectrum of physical, mental, and emotional impairment. Current data on incidence and prevalence of maltreatment in children with disabilities are limited by varying definitions of disability and lack of uniform methods of classifying maltreatment. Nonetheless, children with disabilities and special health care needs are at increased risk of child maltreatment. This report is an update to the previous policy statement, “Assessment of Maltreatment of Children With Disabilities.”

The Children’s Bureau reported that an estimated 872,000 children were determined to be victims of abuse or neglect in 2004. More than 60% of child victims experienced neglect, almost 20% were physically abused, and 10% were sexually abused. Of the 36 states that reported on disabilities, child victims who were reported with a disability accounted for 7.3% of all victims. Children with the following conditions were considered as having a disability: mental retardation, emotional disturbance, visual impairment, learning disability, physical disability, behavioral problems, or another medical problem. It was believed that these conditions were underrecognized and underreported, because not every child received a clinical diagnostic assessment when child maltreatment was suspected.

Child maltreatment may result in the development of disabilities, which in turn can precipitate further abuse. Studies have been unable to accurately document the extent or rate of abuse among children with disabilities or determine if disabilities were present before the abuse or were the direct result of maltreatment. It should be emphasized also that several case reports and epidemiologic data remind us that the natural history of some medical conditions can include conditions that mimic child maltreatment.
The numbers of children who survive disabling medical conditions as a result of technologic advances and children who are recognized and identified as having disabilities are increasing. The rates of child maltreatment have been found to be high in the child population in general and in children with blindness, deafness, chronic illness, developmental delays, behavioral or emotional disorders, and multiple disabilities. Minimal research on child abuse has focused specifically on children with disabilities; further study is indicated and has been encouraged.

The Child Abuse and Prevention, Adoption, and Family Services Act of 1988 (Pub L No. 100–294) included a mandate to study the incidence of child maltreatment among children with disabilities. This research was funded by the National Center on Child Abuse and Neglect and conducted by the Center for Abused Children With Disabilities at the Boys Town National Research Center. A study by Westat, Inc, determined the incidence of abuse among children with disabilities and the relationship between child abuse and disabilities. Data were collected from 35 child protective services (CPS) agencies across the country, and results indicated that 14.1% of children whose maltreatment was substantiated by CPS workers had 1 or more disabilities. A Nebraska study that used an electronic merger of hospital, central registry, foster care review board, and law enforcement records found disabilities to be twice as prevalent among maltreated children in hospitals as among hospital controls, which is consistent with the hypothesis that disabilities increase the risk of maltreatment. The data are also consistent with the hypothesis that maltreatment contributes to disabilities.

According to research performed by the Boys Town National Research Hospital, children with disabilities were found to be at greater risk of becoming victims of abuse and neglect than were children without disabilities. The study showed that children with disabilities are 1.8 times more likely to be neglected, 1.6 times more likely to be physically abused, and 2.2 times more likely to be sexually abused than are children without disabilities. Another study found the overall incidence of child maltreatment to be 39% in 150 children with multiple disabilities admitted to a psychiatric hospital. Of those children, 60% had been physically abused, 45% had been neglected, and 36% had been sexually abused. In a 2000 study of more than 4500 maltreated children, Sullivan and Knutson observed children with disabilities to be 3.76 times more likely to be neglected, 3.79 times more likely to be physically abused, and 3.14 times more likely to be sexually abused when compared with children without disabilities. Children with behavioral disorders were found to be at the highest risk of all types of maltreatment, and neglect was the most common form of maltreatment across all disability types. A relative-risk matrix for all types of maltreatment among children with specific disabilities was developed. In 1 recent study, caregivers reported that 18.5% of children with autism had been physically abused and 16.6% had been sexually abused. Spencer et al concluded that children with disabling conditions are at increased risk of child abuse and neglect, although the type of maltreatment varies with the specific disabling condition.

**LIMITATIONS OF CURRENT RESEARCH**

The prevalence of maltreatment of children with disabilities is difficult to calculate, because states do not use comparable definitions of child abuse and neglect. Another major problem with the published literature is the variable definition of “disabilities.” The Centers for Disease Control and Prevention describes developmental disabilities as a diverse group of severe chronic conditions that are attributable to mental and/or physical impairments and result in problems with major life activities such as language, mobility, learning, self-help, and independent living. The Americans With Disabilities Act defines “disability” as a physical or mental impairment that substantially limits 1 or more of the major life activities of an individual. This definition includes all types of disabilities, including physical disabilities, cognitive or learning disabilities, motor and sensory dysfunctions, mental illness, or any other kind of physical, mental, or emotional impairment. Perrin reported that most childhood chronic health conditions do not cause disability. The Maternal and Child Health Bureau has defined children with special health care needs as “those who have or are at increased risk for a chronic physical, developmental, behavioral, or emotional condition and who also require health and related services of a type or amount beyond that required by children generally.” The term “children with special health care needs” is less limiting than some other terms.

Legal definitions do not always match clinical data. Child development evaluations do not always allow an immediate and precise diagnosis of extent or type of disability, and some studies rely on evaluations by untrained observers. Therefore, evaluation of research efforts is hindered by different definitions of terms (eg, disabilities and maltreatment), noncomparable methods, various study sample sizes, and lack of uniform data collection. Furthermore, changes in reporting laws and societal attitudes can occur during a study period.

Another problem that has been cited in the literature is the lack of recognition and documentation of disabilities by CPS workers and their lack of training on evaluating children with disabilities. In the study by Westat, analyses were based on CPS workers’ opinions rather than data empirically derived from physicians or other professionals trained to diagnose disabilities. Bonner et al demonstrated that since 1982, correct and consistent use of a CPS-created system of collecting information regarding disabilities in maltreated children
had decreased, suggesting that disabilities were unlikely to be identified as children enter the CPS system. A survey of 51 state CPS agencies found that in 86% of states, CPS workers used a standardized form to record child-maltreatment cases, but in only 59% of those states did the workers record information regarding pre-existing disabilities on the form.22

The Westat study was limited to intrafamilial cases.9 Because it is well known that individuals other than family members can commit harm to children, statistics limited to intrafamilial cases would be likely to underestimate the overall incidence of maltreatment among children with disabilities.

Along with the lack of well-designed research on maltreatment is the lack of research on how to respond to children with disabilities who have been maltreated. A needs assessment of parents, educators, and CPS investigators in 2000 revealed that knowledge and experience in child maltreatment were lacking.23 Most respondents were interested in training, with recognition of maltreatment in the child with a disability as a top priority. In this same report, a collaborative team-approach response was recommended.

CAUSAL FACTORS

In general, the causes of abuse and neglect of children with disabilities are the same as those for all children; however, several elements may increase the risk of abuse for children with disabilities. Children with chronic illnesses or disabilities often place higher emotional, physical, economic, and social demands on their families.21 For example, a physical disability that causes difficulty in ambulation can place a child at risk of accidental falls. Therefore, much closer supervision will be needed, which itself can be stressful. Parents with limited social and community support may be at especially high risk of maltreating children with disabilities, because they may feel more overwhelmed and unable to cope with the care and supervision responsibilities that are required.17 Lack of respite or breaks in child care responsibilities can contribute to an increased risk of abuse and neglect. Finally, the added requirements of special health care and educational needs can result in failure of the child to receive needed medications, adequate medical care, and appropriate educational placements, resulting in child neglect.17

Numerous problems have been cited with the provision of care for foster children with disabilities. Foster parents sometimes are not told about a child’s medical and emotional problems and are, therefore, not sufficiently educated or prepared to deal with the specific condition. Other problems for foster children with disabilities include lack of permanent placement, lack of a medical home, lack of financial support, and inappropriately prepared foster parents.3

Parents or caregivers may feel increased stress because children with disabilities may not respond to traditional means of reinforcement, and sometimes these children’s behavioral characteristics (eg, aggressiveness, noncompliance, and communication problems) may become quite frustrating.2 Behavioral challenges in children who have disabilities may further increase the likelihood of physical abuse.17

Parents of children with communication problems may resort to physical discipline because of frustration over what they perceive as intentional failure to respond to verbal guidance. It has been noted paradoxically, however, that families who report higher stress levels may actually have greater insight into problems associated with caring for a disabled child, whereas parents with a history of neglect of a child may not experience the level of stress that a more involved parent may experience.24

Although the use of aversive procedures and restraints for children who have disabilities has been fortunately diminishing, in part because of legislative changes (eg, modifications of the Individuals With Disabilities Education Act [Pub L No. 108–446 (2004)]), these practices are still used sometimes in homes, schools, programs, and institutions.25 Aversive techniques are procedures that use painful or unpleasant stimuli to modify a behavior that has been found to be unacceptable or inappropriate. Restraints are physical measures (such as tie-downs or prolonged seclusion) used to prevent something physical from happening or for “punishment.” This includes “therapeutic holding,” which has been repudiated as being harmful.26 During the past 20 years, much research has demonstrated the effectiveness of alternative measures, commonly called “positive behavioral supports,” to change behavior.25 Pediatricians and others who could use additional information about the problems that occur from the use of aversive procedures or restraints can easily get this guidance from the Web sites of organizations such as the Association for the Severely Handicapped (www.tash.org/IRR/resolutions/res02aversive.htm) and the Autism National Committee (www.autcom.org/restraints.html). Information about positive behavioral support guidelines is available from a US Department of Education–funded program, the Technical Assistance Center on Positive Behavioral Interventions & Supports (www.pbis.org), and other national and international programs. The American Academy of Child and Adolescent Psychiatry also provides guidance on this subject (www.aacap.org/page.ww?section=Policy+Statements&name=Coercive+Interventions+for+Reactive+Attachment+Disorder).

The presence of multiple caregivers may heighten or reduce the risk of abuse of the child. Infrequent contact of a child with disabilities with other children and adults may make them uniquely vulnerable to molestation because there is decreased opportunity for the child to
develop a trusting relationship with an individual to whom he or she may disclose the abuse and decreased opportunity to learn to resist molestation. On the other hand, children with disabilities who require multiple caregivers or providers may have contact with numerous individuals, thereby increasing the opportunity for abuse, including sexual abuse. However, advantages to having a large number of caregivers are that there are more individuals who may detect the injuries or signs of abuse, and the additional assistance may actually lessen the amount of stress placed on the primary caregivers. Risk may be minimized by careful screening and selection of caregivers, sporadic and unscheduled monitoring of care, and an open mind to recognition that any child may become a victim.

Children with disabilities often have limited access to critical information pertaining to personal safety and sexual abuse prevention. Children who have increased dependency on caregivers for their physical needs may be accustomed to having their bodies touched by adults on a regular basis. Parents may object to their child being provided with education on human sexuality, because they may feel that their children will never be in sexually risky situations because of their special needs. However, children with disabilities may be unintentionally conditioned to comply with authority, which could result in them failing to recognize abusive behaviors as maltreatment. Children with disabilities are often perceived as easy targets, because their intellectual limitations may prevent them from being able to discern the experience as abuse and impaired communication abilities may prevent them from disclosing abuse. Because some forms of therapy may be painful (eg, injections or manipulation as part of physical therapy), the child may not be able to differentiate “appropriate” pain from “inappropriate” pain.

**PEDIATRICIAN’S ROLE**

Pediatricians should be aware that the presence of disabilities in a child is a risk factor for victimization and that disabilities can also be the result of child maltreatment. The pediatrician should work with families, other health care professionals, and other community resources to ensure the safety of all children, including those with disabilities. The following should be considered.

**Identification and Reporting**

Pediatricians always need be alert to signs or symptoms that are suggestive of abuse, no less in children with disabilities than in others. However, recognizing the signs and symptoms of maltreatment among children with disabilities may be difficult, because many children may not be able to verbalize that they were abused or they may not understand that what took place was wrong. Because it is common for the physical examination to be normal, especially in sexual abuse and emotional maltreatment, a high index of suspicion in selected cases is warranted.

Familiarity with the natural history of disorders that may mimic child abuse can prevent the misdiagnosis of child maltreatment. Children with motor and balance disabilities may experience increased injuries from accidents. Children with neurosensory disabilities may be predisposed to fractures, and in the absence of pain, there may be a delay in seeking medical attention. For example, children with spina bifida have a high risk of fracturing a paralyzed, desensitized limb. Children with severe nutritional deficiency and immobility or chronic steroid use may develop demineralized bones that fracture easily. Children with blood dyscrasias may have bruises of varying ages in unusual places. There are also reports of a variety of disabling conditions that mimic child maltreatment, including methylmalonic aciduria and glutaric aciduria, which can manifest as chronic subdural effusions and mild retinal hemorrhages, with other telltale findings including neurodevelopmental problems.

Awareness of injury patterns from inflicted versus noninflicted trauma is important for pediatricians and other professionals who work with children. Signs and symptoms of maltreatment in children with disabilities are commonly ignored, misinterpreted, or misunderstood. Furthermore, many schools, programs, and institutions may have a disincentive to recognize or report child maltreatment because of fear of negative publicity or loss of funding or licensure. Pediatricians may want to act proactively with these entities so that concerns and referrals are more forthcoming if questions or problems arise.

If abuse or neglect is suspected after a careful assessment, a report must be made to the appropriate CPS agency. Every child suspected of being abused or neglected needs a thorough evaluation by an experienced professional. The evaluation process should consist of a structured interview with the child, if possible, and a comprehensive physical examination including appropriate laboratory and radiologic studies as indicated. In many situations, a consultation with a developmental pediatrician, pediatric neurologist, child abuse pediatrician, or other expert in children with disabilities is also indicated.

**Treatment**

Appropriate medical treatment for injuries, infections, or other conditions should be provided. Each case of abuse or neglect that is clinically confirmed or strongly suspected needs a multidisciplinary treatment plan, which includes a mental health assessment and treatment component that is appropriate for the child’s cognitive and developmental level and counseling for the family. This child and family treatment plan should be integrated...
with other intervention plans that may already have been developed for the child. Federal legislation requires that each child identified as having a disability should have a written plan of service (an individualized family service plan for children from birth to 3 years of age or an individualized education plan for children 3 through 21 years of age). Third, the pediatrician should also make appropriate medical recommendations and provide treatments that are preventive or prescriptive. The pediatrician may help the family explore available child care and respite services. A discussion of injury-prevention guidelines for children with disabilities is also helpful. Although pediatricians can have input into the process, removal of the child from the home or therapeutic foster care placement is at the discretion of the CPS agency only after a thorough investigation.

**Education**

In-service training for CPS workers, law enforcement professionals, health care professionals, child care providers, early childhood educators, teachers, and judges is crucial, and protocols are necessary for the identification, reporting, and referral of all cases of suspected child maltreatment in all schools, programs, and institutional settings. Experts in child maltreatment and childhood disabilities would be the main group to assist with this educational endeavor. However, general pediatricians can have important roles, for example, with local school districts. In addition, education on risk factors for maltreatment of children with disabilities should be emphasized.

Pediatricians can also be important role models to parents, trainees, and others. In their own offices, clinics, or hospital settings, pediatricians and others who provide care for children with disabilities should not rush to use physical restraints during procedures for these children. Often, taking the time to explain procedures in terms appropriate to developmental level or in other ways to prepare such a child can make restraints unnecessary except in situations when children are dangerous to themselves or others. Even when some types of restraints are needed, such as to prevent a child from scratching at newly repaired lacerations, such restraints should be as comfortable and as minimal as possible and used for the shortest time feasible.

Pediatricians may also have roles to assist in the education about child abuse to their peers, residents, medical students, and other health care students. All health care professionals need adequate training to monitor children with disabilities for signs of abuse and neglect and to screen suspected victims of child maltreatment for disabilities.

**Prevention**

Support and assistance with parenting skills are often needed by families, especially families who have children with special health care needs. Medical and non-medical needs of the child and family should be addressed at each health supervision visit. Child and family strengths should be recognized and fostered at each encounter. Family stressors should be identified and addressed, and referrals for appropriate support services should be made. Disability-specific injury-prevention guidelines can be presented to help the family minimize injury. The availability of parent support groups, respite care, and home health services may be explored, and referrals may be made when appropriate. Pediatricians can help educate parents of children with disabilities about the various respite and medical waiver subsidies and programs specific to each state and how to qualify for such funds. Pediatricians can explain the need for getting placed on the inevitable waiting lists for these programs as early as possible. The American Academy of Pediatrics (AAP)–sponsored medical home Web site (www.medicalhomeinfo.org) is an important resource for the pediatrician to find out more information on these programs, including state-by-state resources.

All children with or without disabilities need a medical home that consists of a health care professional who is readily accessible to the family to answer questions, help coordinate care, and discuss concerns. Medical home is an approach to providing comprehensive primary care that is accessible, continuous, comprehensive, family centered, coordinated, compassionate, and culturally effective. Families should be encouraged and assisted by these health care professionals to work with a variety of agencies and disciplines and pursue resources and services that they need. Child abuse prevention, including indicators of abuse, should be discussed with parents and caregivers. Neurodevelopmental and developmental-behavioral pediatricians and child neurologists who are trained and experienced in the diagnosis and evaluation of children with disabilities can also serve as excellent resources to both the general pediatricians and the families.

**Advocacy**

The pediatrician, in providing the medical home and acting as his or her patient’s and family’s advocate, may review care that is provided by the various agencies and resources. Much of this advocacy effort can be performed by coordinating efforts and ensuring that recommendations are made and followed. By providing this careful follow-up, if child maltreatment is suspected, the need for appropriate referrals can be identified immediately.

Pediatricians play an important advocacy role in their relationships with various governmental and nongovernmental agencies. State AAP chapters also have an important role in these arenas. State, educational, social, foster care, financial, and health care systems often function in isolation from each other, with very little coordination or communication.
can also encourage the development of needed resources, including child care and respite services for families with a child with a special health care need. Foster children with disabilities and their foster parents often suffer from lack of adequate support systems. Communication with schools and other systems with which families interact is another avenue for heightening the awareness of the needs of children who have disabilities and/or special health care needs.

As child advocates, pediatricians are in an ideal position to influence public policy by sharing information and giving educational presentations on child maltreatment and the needs of children with disabilities. They should advocate for state practices or policies that mandate CPS agencies to gather disability information on child-maltreatment cases. Pediatricians should emphasize the devastating costs of child maltreatment to legislators, policy-makers, and the public. Pediatricians should also advocate for screening procedures for potential employees in educational, recreational, and residential settings to help ensure the safety of all children in their care.

One resource that is useful to the pediatrician is the report *A Call to Action: Ending Crimes of Violence Against Children and Adults With Disabilities.* This is a report that includes recommendations on policy, surveillance systems and data collection, violence prevention, intervention, and research needs. The Oregon Institute on Disability & Development has developed prevention resources that may be useful to the pediatrician (www.ohsu.edu/research/oidd/oakspublication.cfm).

**GUIDANCE FOR THE PEDIATRICIAN**

1. Be capable of recognizing signs and symptoms of child maltreatment in all children and adolescents, including those with disabilities.

2. Be familiar with disabling conditions that can mimic abuse or pose an increased risk of accidental injury that can be confused with abuse.

3. Because children with disabilities are at increased risk of maltreatment, remain vigilant not only in assessment for indications of abuse but also in offer-ings of emotional support and provision of equipment and resources to meet the needs of children and families.

4. Ensure that any child in whom maltreatment has been identified is evaluated thoroughly for disabilities.

5. Advocate for all children, especially those who have disabilities or special health care needs, to have a medical home. If a child is hospitalized and does not have a medical home, the inpatient attending physician can help the family secure one before discharge, preferably as early as possible in the hospital course.

6. Be actively involved with treatment plans developed for children with disabilities and participate in collaborative team approaches.

7. Use health supervision visits as a time to assess a family’s strengths and need for resources to counterbalance family stressors and parenting demands.

8. Advocate for changes in state and local policies in which system failures seem to occur regarding the identification, treatment, and prevention of maltreatment of children with disabilities.

9. Advocate for the implementation of positive behavioral supports and elimination of aversive techniques and unnecessary physical restraints in homes, schools, and other educational and therapeutic programs (both public and private), institutions, and settings for children who have disabilities.

10. Advocate for better health care coverage by both private insurers and governmental funding.

**CONCLUSIONS**

The AAP supports the belief that pediatricians play a significant role in the prevention, identification, and treatment of child abuse and neglect, especially in children with disabilities, who are at increased risk of maltreatment. Children suspected of being maltreated should be evaluated for developmental, physical, and mental health disabilities. In addition, CPS workers and others involved in the investigation of child maltreatment should work closely with pediatricians to identify disabilities in children. Every effort should be made to ensure the safety of children through collaboration with families, other health care professionals, schools, CPS agencies, and other appropriate resources.

**COMMITTEE ON CHILD ABUSE AND NEGLECT, 2006–2007**

Carole Jenny, MD, MBA, Chairperson
Cindy W. Christian, MD
*Roberta A. Hibbard, MD
Nancy D. Kellogg, MD
Betty S. Spivack, MD
John Stirling, Jr, MD

**liaisons**

David L. Corwin, MD
American Academy of Child and Adolescent Psychiatry
James A. Mercy, PhD
Centers for Disease Control and Prevention

**Staff**

Tammy Hurley
REFERENCES


The rise of in-store clinics: threat or opportunity?

“The recent acquisition by the pharmacy chain CVS of MinuteClinic, a chain of in-store clinics founded in Minnesota, has put this model of primary care delivery back in the spotlight. Although still not widespread, the model is increasing in prevalence . . . and appeals to several stakeholders: payers note that primary care is less expensive when delivered at in-store clinics than when provided in a doctor’s office or emergency room, patients value the convenience and low price, entrepreneurs see a profitable business model, and proponents of consumer-driven health care see services that can be paid for out of health savings accounts. Physicians, however, express concern about the quality of care and the potential impact on their businesses. The typical in-store clinic is a kiosk—a small, thin-walled structure located inside a store—staffed by a nurse practitioner. The clinics differ from the old ‘doc-in-the-box’ model in that they are neither routinely staffed by a physician nor intended to provide all primary care services. Indeed, the range of services—posted as a ‘menu’ on the company’s Web site or on the kiosk—is strikingly small, including common adult vaccinations, screening tests, and treatment for simple conditions. . . . But for these circumscribed services, the clinics provide a compelling value proposition. Care is intended to be quick, inexpensive, and convenient; visits and waiting times are short, the charge is usually less than $50, and extended hours are offered along with ample parking. It’s not surprising, then, that patients and investors have taken notice. . . . Some wonder whether this model is a ‘disruptive innovation’—that is, a service or technology that enters a market at the low end, initially not performing as well as higher-end incumbents, then improves until it captures the whole market.”


Noted by JFL, MD
Beyond Munchausen Syndrome by Proxy: Identification and Treatment of Child Abuse in a Medical Setting

John Stirling, Jr, MD, and the Committee on Child Abuse and Neglect

ABSTRACT
The condition widely known as Munchausen syndrome by proxy comprises both physical abuse and medical neglect and is also a form of psychological maltreatment. Although it is a relatively rare form of child abuse, pediatricians need to have a high index of suspicion when faced with seemingly inexplicable findings or treatment failures. The fabrication of a pediatric illness is a form of child abuse and not merely a mental health disorder, and there is a possibility of an extremely poor prognosis if the child is left in the home. In this statement, factors are identified that may help the physician recognize this insidious type of child abuse that occurs in a medical setting, and recommendations are provided for physicians regarding when to report a case to their state’s child protective service agency.

INTRODUCTION
In the oft-quoted paraphrase of Hippocrates, the physician is admonished to “first, do no harm,” and not without good reason. Even when necessary, diagnostic tests are at best inconvenient and frequently invasive or painful. Therapy is not without risk either, because it often involves hospitalization, drugs, or surgery. When the diagnosis is elusive and diagnostic efforts become more aggressive, the physician must always weigh risks to the patient against the benefits of an accurate diagnosis. Nowhere does this calculation become more important than in the rare circumstance in which the patient’s caregiver fabricates the signs or symptoms of the disease in question, in what has traditionally been called Munchausen syndrome by proxy.

DESCRIPTION
The fictitious Baron von Munchausen was an extravagant raconteur, whose fanciful narrations of his imagined exploits made his name in literature. Physicians have borrowed his name to describe a group of patients whose complaints are fabricated but so convincing that they are subjected to needless hospitalizations, laboratory tests, and even surgery. The *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV) refers to Munchausen syndrome as “factitious disorder” (300.19), and motivations for this bizarre behavior continue to puzzle both medical and mental health professionals.

In 1977, Meadow first described cases in which the apparent symptoms of Munchausen syndrome were instead projected onto a dependent child as a parent fabricated symptoms and even signs of a nonexistent illness. When the fabrica-
tions involved a dependent individual like this, the condition was likened to Munchausen syndrome experienced “by proxy,” and the diagnosis of Munchausen syndrome by proxy entered the medical lexicon. In the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, the condition is proposed as a new category called “factitious disorder by proxy.”

There is no typical presentation for this condition. Suspicions may arise when parents misinterpret or exaggerate normal behaviors, and true cases range from apparent fabrication of reported symptoms to outright fabrication of signs of disease. Caregivers may report signs and symptoms that are undetectable to the medical observer, or the child may demonstrate signs that defy medical interpretation. In case reports, a wide variety of situations have been called, appropriately or inappropriately, Munchausen syndrome by proxy, including the following examples:

- A mother takes her child to the doctor for frequent evaluations for sexual abuse, even in the absence of objective evidence or history of abuse.5
- Mothers insist their children be treated for attention-deficit/hyperactivity disorder although there is no evidence to make the diagnosis.6
- A parent starves her child because she wrongly believes he has multiple food allergies.7
- Physicians suspect an unusual hematologic disorder after a mother repeatedly and secretly bruises her child with a hammer.8
- A parent purposely suffocates her child and kills him during a hospitalization for “apnea.”9

It is difficult to imagine how such varied conditions can be included in the definition of a syndrome. In some cases, the caregiver has merely exaggerated the child’s symptoms; in others, the caregiver has imagined them. In the worst cases, the signs and symptoms of illness have been induced by the caregiver’s intentional actions. In some patients, the consequences are minor; in others, the consequences are fatal. Indeed, the only things common to the presentations cataloged above are the caregivers’ insistence that something was wrong, an absence of pathologic findings sufficient to explain the described signs or symptoms, and consequent harm to the child.

**TERMINOLOGY**

Use of the term “Munchausen syndrome by proxy” has led to much confusion in the literature. For example, some experts insist that the term be applied only when the parent is seeking medical care because they are somehow personally compelled to relate to the medical care system,10,11 whereas others say the parent’s motivation is not important.5,12 Although the original description referred to harmful medical care, subsequent authors have extended the appellation “Munchausen syndrome by proxy” to cases in which the only harm arose from medical neglect or noncompliance13,14 or even educational interference.9 In addition, there remains confusion about who should make the diagnosis of Munchausen syndrome by proxy: a psychiatrist or pediatrician? Is it a diagnosis applied to the parent or the child? Is it a pediatric or a mental health diagnosis? These ambiguities become especially important when medical personnel present their diagnosis to other professionals or to juries in seeking to protect a child victim.

To alleviate confusion, the American Professional Society on the Abuse of Children has recently made a more explicit distinction between the abuse (pediatric condition falsification) and the presumed motive behind most such cases (factitious disorder by proxy).1 This distinction has the advantage of replacing an eponym with more descriptive nomenclature, a recent and welcome trend in medicine. Whatever it is called, it is important to remember that harm incurred when a caregiver exaggerates, fabricates, or induces symptoms of a medical condition may still simply be termed “child abuse, which happens to occur in a medical setting.” This appellation reminds us that the focus of our intervention should always be to identify and minimize harm to the child regardless of the motivation of the perpetrator.

**DEFINITION**

Whether it is called Munchausen syndrome by proxy, pediatric symptom falsification, or simply child abuse, what remains as the central issue of importance is that a caregiver causes injury to a child that involves unnecessary and harmful or potentially harmful medical care. To make the diagnosis, the physician must ask 3 questions:

1. Are the history, signs, and symptoms of disease credible?
2. Is the child receiving unnecessary and harmful or potentially harmful medical care?
3. If so, who is instigating the evaluations and treatment?

If the child receives excessive, unnecessary medical care merely because the physician is overly compulsive or, worse, incompetent, then abuse is not a consideration. If the child is getting the unnecessary medical care because the parent is systematically misrepresenting symptoms, purposefully making up symptoms, manipulating laboratory tests, or even purposefully harming the child to create symptoms (eg, by poisoning or suffocation), then continued medical care itself may become abusive. The medical staff, in pursuing an ever-more-elusive organic diagnosis, may lose sight of its ultimate implausibility. One needs 2 circumstances to make the diagnosis in this form of abuse: harm or potential harm.
to the child involving medical care and a caregiver who is causing it to happen.

The motive of the caregiver, although useful to the therapist, is unimportant in making the diagnosis of abuse. In no other form of child abuse do we include the perpetrator’s motives as a diagnostic criterion. For example, a man can sexually abuse a child for a variety of reasons, but his motivation is irrelevant; the child still carries the diagnosis of sexual child abuse. A mother might violently physically assault her infant because she is fed up with the child crying, she is intoxicated or drugged, or she earnestly thinks that is the way to get the infant to behave and start eating, but it is still called physical child abuse.

Child abuse is a pediatric diagnosis, one that describes what is happening to the child. Motivation of the perpetrator often becomes an issue when society considers incarceration, treatment, or reunification but not when the physician makes the medical diagnosis of child abuse.

**DIAGNOSIS**

Diagnosis of fabricated disease can be especially difficult, because the signs and symptoms are undetectable (when they are being exaggerated or imagined) or inconsistent (when they are induced or fabricated). Researchers may differentiate between exaggeration and fabrication or induction of symptoms, but action taken by the clinician must be determined by the perception of harm or potential harm to the child.

Regardless of the exact nature of the duplicity, health care professionals can be seduced into prescribing diagnostic tests and therapies that are potentially injurious. This is easier than one might think. After all, absolute certainty is a rare thing in medical diagnosis, and physicians have all known empirical therapy to be effective. On occasion, though, the well-meaning but misguided pursuit of an ever-more-elusive diagnosis or effective treatment can lead medical staff into an ethical dilemma. Potentially harmful medical care can range from a diagnostic search that subtly encourages and enables a caregiver’s delusion through a full spectrum of invasive tests and medical or even surgical interventions. Alternatively, a child may present to the doctor with a common diagnosis but one that seems resistant to an increasingly aggressive array of treatment regimens. The common factor in all is the failure to consider factitious disease in the differential diagnosis, although it is often more likely than the arcane diagnoses being pursued so assiduously.

Child abuse is not a diagnosis of exclusion. On the contrary, when a clinician suspects that a disease has been falsified, this hypothesis must be pursued vigorously and the diagnosis must be confirmed if the child is to be spared further harm. In seeking to determine if signs and symptoms of a disease have been fabricated, the physician should make every effort to gather information from all those involved and make other professionals aware of the concerns. Care of children who are victims of factitious disorder by proxy often involves a variety of medical personnel, from primary care physicians and medical subspecialty consultants to dietitians, physical therapists, and social service workers, and each has a unique perspective. Nursing and support staff can frequently contribute to making the correct diagnosis by reporting their observations of, and experiences with, the child and family to the supervising physician. It should be stressed, however, that the falsification of a medical condition is a medical diagnosis. Although multidisciplinary input can be very helpful in diagnosis and essential in treatment, psychologists, social workers, and others are not in a position to make or confirm this diagnosis.

Occasionally, more information about the maltreatment is needed before a diagnosis can be reached. When it is suspected that no true disease exists and it is felt that harm to the child is imminent, the use of covert videotape surveillance has been recommended. Such surveillance may capture a parent’s misbehavior, as when a child is being physically abused in the hospital. It may fail to confirm reported symptoms when they are being exaggerated or exonerate a suspected caregiver when a disease truly exists. In any event, video surveillance cannot be considered a gold standard or held as the only way of diagnosing this insidious form of child abuse. When videotaping is used, adequate safeguards such as continuous surveillance and a well-understood plan of action must be present to prevent further injury.

**TREATMENT**

By recognizing that this problem is a form of child abuse taking place in a medical setting, a clear role is delineated for the system that is currently in place in our states to protect children. Child protective services agencies are mandated to keep children who are abused—sexually, physically, or psychologically—safe regardless of whether the abuse occurs in the home or the hospital.

When considering treatment for child abuse taking place in a medical setting, the basic principles used in any other type of child abuse case should be applied:

1. Make sure the child is safe.
2. Make sure the child’s future safety is also assured.
3. Allow treatment to occur in the least restrictive setting possible.

For example, if an overanxious mother who has insisted on too much medical care for her child is willing to cooperate with the physician and learn when it is appropriate to seek care, the child can safely be treated within his or her family setting. In contrast, if a mother has repeatedly suffocated her child, the “least restrictive
setting” that would guarantee the child’s safety would most likely be permanent out-of-home placement.

If the parent’s care-seeking is harming the child but the parent refuses to cooperate with the physician in limiting the amount of medical care to an appropriate level, the state child protective services agency should be informed. If the parent persists in harming the child, medical child abuse should be reported in the same way as physical and sexual child abuse. Any time that a dependent child is being hurt by an adult’s action, child protective services should become involved.

A list of possible interventions follows, from the least restrictive to the most restrictive. Some of these options require action by outside agencies (child protective services, private counselors, law enforcement, etc).

1. Use individual and/or family therapy while depending on a primary care physician to be “gatekeeper” for future medical care utilization.

2. Monitor ongoing medical care usage by involving people or institutions outside the medical practice to alert the physician gatekeeper about health care issues. For example, in the event of a child protective services investigation, or with the parent’s consent, the insurance provider can be alerted to inform the primary care physician or medical home about visits to other professionals. Another example would be having the parent authorize the school to call the physician any time the child is absent or have school officials agree not to excuse any absence without the physician’s approval.

3. Admit the child to an inpatient hospital setting or a partial hospital program, where his or her actual signs and symptoms can be monitored (as opposed to the signs and symptoms reported by the parent). This admission is a very important resource if the parent tends to exaggerate or lie about the child’s pain or disability. A program that treats the whole family can then work to define the child as normal in the parents’ eyes.

4. Involve child protective services to obtain dependency, either in or out of the home, to control overuse of medical resources and gradually reintroduce the child to the caregiver’s home while monitoring the child’s safety.

5. Place the child in another family setting permanently.

6. Prosecute the offending parent and incarcerate him or her, thus eliminating access to the child.

The physician’s role in options 4 through 6 would be to report the case to the appropriate authorities, carefully document the abuse, and, if needed, testify on the child’s behalf in courts of law. Obviously, options 3 through 6 will be required only in the most extreme or persistent cases of medical abuse.

CLINICAL ADVICE

When physicians diagnose and manage cases of child abuse in the medical setting, the following clinical advice will help ensure a more successful outcome of the case:

1. Whenever possible, have a pediatrician with experience and expertise in child abuse consult on the case, if not lead the team. This may help to reduce “false-positive” misdiagnosis and better identify actual cases.

2. Review all the medical charts pertinent to these complicated cases. Abusing parents often seek medical care from a variety of sources and may change physicians frequently. It is important to involve all the treating physicians in the process. Primary care and subspecialty physicians should work together to identify parents who seek excessive medical care. They should communicate regularly about the degree of medical care utilization and reach consensus on management. Cooperation of all the involved physicians is not only critical to good patient care, but it can also keep the parent from becoming confused or deliberately playing one doctor against another.

3. Work with a hospital- or community-based multidisciplinary child protection team. Such teams bring a variety of skills and viewpoints to the treatment process and provide expert consultation for the primary care physician in child maltreatment and child protection.

4. When a “more restrictive” response is needed, do not hesitate to involve the state social service agency responsible for protecting children from abuse. If the physician has access to a multidisciplinary child protection team, the team can help coordinate efforts to protect the child and facilitate communication with the state child protection agency.

5. Involve the whole family in the treatment. Their entire view of illness and health in their lives has to be adjusted. Ongoing family issues must be addressed to guarantee the future safety of the victim and any other children in the home. Therapists may use effective behavioral management techniques to change the child’s dysfunctional behaviors, when appropriate.

SUMMARY

What has been known as Munchausen syndrome by proxy may be better described as pediatric condition falsification or simply child abuse that occurs in a medical setting. In aggressively seeking an elusive diagnosis, physicians can sometimes cause harm to their patient and must remain aware of this possibility. The pediatrician who suspects that signs or symptoms of a disease are in fact being fabricated should concentrate on the harm or potential harm to the child caused by the actions of that caregiver and the efforts of the medical personnel to
diagnose and treat a nonexistent disease. Proper diagnosis of fabricated disease involves thorough evaluation of medical charts, clear communication among medical professionals, and, often, a multidisciplinary approach. A focus on the motives of the caregiver, although useful in therapy, is unnecessary for the diagnosis of this form of child abuse.

COMMITTEE ON CHILD ABUSE AND NEGLECT, 2006–2007
Carole Jenny, MD, MBA, Chairperson
Cindy Christian, MD
Robert A. Hibbard, MD
Nancy D. Kellogg, MD
Betty S. Spivak, MD
John Stirling, Jr, MD

LIAISONS
David L. Corwin, MD
American Academy of Child and Adolescent Psychiatry

STAFF
Tammy Piazza Hurley

REFERENCES
AAP Publications Reaffirmed,
January 2007


Movement of Bilirubin and Bilirubin Conjugates Across the Placenta

To the Editor.—

The recent report of a woman with conjugated hyperbilirubinemia who delivered an infant with negligible concentrations of bilirubin in serum at birth echoes an earlier report of a similar phenomenon. In 1973, Lipsitz et al described a pregnant patient with a total plasma bilirubin concentration of 24 mg% (13 mg% direct bilirubin) on the day of delivery and whose male infant showed a cord blood total bilirubin concentration of 10.4 mg%, of which only 0.6 mg% was direct-reacting bilirubin. These observations are consistent with experimental studies in rhesus monkeys that showed that conjugated bilirubin does not cross the placenta from the maternal to the fetal circulation. Investigations in the guinea pig and monkey also showed that bilirubin glucuronides do not cross the placenta in the reverse direction, from fetus to mother. Thus, bilirubin glucuronides do not cross the placenta readily in either direction in vivo.

In the patients noted above, the direct-reacting pigment in maternal plasma would have been a mixture containing both bilirubin glucuronides and isomers of these resulting from acyl migration of the glucuronic acid, all of which would give the diazo reaction for direct bilirubin. Therefore, not only bilirubin glucuronides but other glucuronic acid esters of bilirubin fail to cross the human placenta readily. Although the placenta contains membrane transport proteins that effect the efflux of bilirubin glucuronides from the liver, it would seem that these transporters do not facilitate placentation passage of bilirubin glucuronides or their acyl-migrated isomers in either direction in pregnant women.

The observations of mothers with conjugated hyperbilirubinemia contrast markedly with observations of mothers with unconjugated hyperbilirubinemia, a situation that is seen in patients with Crigler-Najjar syndrome. Because of a congenital deficiency in bilirubin conjugation, these patients have lifelong unconjugated hyperbilirubinemia. Healthy newborns of mothers with Crigler-Najjar syndrome invariably have elevated plasma levels of unconjugated bilirubin at birth. Notably, the plasma concentrations of bilirubin observed in the newborns are roughly the same as those in the mother. For example, in a recent report, the total bilirubin concentration (all indirect) was 242 μM (14.2 mg%) in the mother at delivery, 247 μM (14.4 mg%) in the newborn, and 222 μM (13.0 mg%) in both the umbilical artery and vein. (Similarly, in mothers with marked elevations of both conjugated and unconjugated bilirubin fractions in plasma, the concentration of unconjugated bilirubin in neonatal blood has been found to be similar to that in the maternal circulation.) These observations are consistent with experiments in the guinea pig and monkey, which have shown that unconjugated bilirubin, in contrast to bilirubin glucuronides, passes readily in both directions across the placenta. They challenge the notion that bilirubin does not readily cross from the maternal to the fetal circulation. Furthermore, the apparent close equivalence between maternal and neonatal (fetal) unconjugated bilirubin concentrations in humans strongly suggests that passive diffusion is the predominant mechanism for the bidirectional placental flux of bilirubin in these patients and that active transport is of little, if any, importance. It is possible, however, that under these pathologic conditions bilirubin saturates placental membrane transport protein(s), which under normal conditions might facilitate efflux of bilirubin from the fetal to the placental circulation.

Antony F. McDonagh, PhD
Division of Gastroenterology
Department of Medicine
University of California
San Francisco, CA 94143
We are grateful to Dr McDonagh for underlining the different transport modalities across the placenta of bilirubin in its conjugated versus unconjugated form. The latter, indeed, seems to be more mobile than the presence of specific carriers would predict. In this respect, we recently showed in vitro and in vivo that unconjugated bilirubin increases the permeability to small molecules across the intestine. Unconjugated bilirubin interacts with the tight junction-protein complex between adjacent enterocytes, “opening the gates” from and to the intestinal lumen. Because these protein structures regulate most biological barriers (including the placenta and the blood-brain barrier), bilirubin may have a carrier-independent, previously unreported way to promote its own transport across the main biological fences.
Children are caught in the middle of an ambiguous relationship between their immigrant parents and the US government.

No one is better qualified to deal with the complexities of this relationship and its impact on child health than pediatricians and other child health care providers. Pediatricians must take their expertise in child health to the next level by advocating for immigration-policy reform. If they do not, children will continue to be an afterthought in immigration debate and policy and will continue to suffer.

Eugenia Garvin, BA
Washington University School of Medicine
St Louis, MO 63110-1093

REFERENCES

In Reply.—

We thank Ms Garvin for her letter regarding our article and commend her early interest as a medical student in advocating for immigrant children and their families.

We agree wholeheartedly that pediatricians can and should play an important role in shaping immigration policy. We take issue, however, with the statement that pediatricians have not traditionally engaged in shaping immigration policy. Indeed, there is clear and abundant evidence to the contrary. For example, several pediatricians recently testified on behalf of the American Academy of Pediatrics about the impact of policies on immigrant children’s health care before several committees of the US House of Representatives, including the House Committee on Government Reform, the House Judiciary Committee, and the House Committee on Energy and Commerce.

Pediatricians have published narrative accounts about immigrant patients who cogently articulate the need for immigrant health care policy reform. A recent article by pediatricians described the case of an undocumented immigrant child in acute renal failure who was refused dialysis by a hospital, and these authors also analyzed the key policy issues for health care providers and policy-makers when faced with caring for undocumented children who need expensive life-saving care. This case and the article were key components of a subsequent challenge by the Mexican American Legal Defense and Educational Fund and a broad coalition of other organizations of the constitutionality of Proposition 200 in Arizona, which cut off all state services, including education, medical care, and police and fire services, to all individuals who are unable to immediately provide adequate proof of their US citizenship or residence.

We also agree, however, that much more policy work is needed on behalf of immigrant children’s health care in the United States. Perhaps most importantly, every child in our nation, regardless of immigration status, should have health insurance. Several states, including California, Oregon, and Washington, are currently considering bills that would provide universal health insurance coverage for children, including all immigrant children and the children of all immigrant parents. We believe that such universal coverage makes sense, will enhance children’s health care access and outcomes, and should be instituted throughout our nation.

We invite Ms Garvin and all pediatricians to join us in advocating on behalf of immigrant children and their families so that no child in America ever suffers, is sick, or dies because of discrimination or lack of access to health care because of immigration status, national origin, race/ethnicity, primary language, or English proficiency.

Glenn Flores, MD
Sandra C. Tomany-Korman, MS
Center for the Advancement of Underserved Children
Department of Pediatrics
Medical College of Wisconsin and Children’s Hospital of Wisconsin
Milwaukee, WI 53226

Milagros Abreu, MD, MPH
Department of Epidemiology
Boston University School of Public Health
Boston, MA 02118

REFERENCES
Is the Binge-Drinking Glass Half Full or Half Empty?

To the Editor.—

The recent article by Miller et al1 reaffirms the risks associated with heavy episodic drinking among adolescents. However disturbing these findings, it is important to recognize some positive trends regarding alcohol use among teens. First, it should be emphasized that 55% of respondents indicated no alcohol consumption in the previous 30 days, and 16.1% had consumed alcohol but not at the binge-drinking level. Thus, the vast majority of teens had actually not engaged in high-risk alcohol consumption in the preceding month. That is very good news.

Most importantly, the authors did not address recent substantial downward trends in heavy episodic drinking among teens. The Youth Risk Behavior Survey results from 1991–2005 demonstrated declines in heavy episodic drinking, and since 1997 the rates have decreased from 33.4% to 25.5%.2 These changes represent rather dramatic shifts among the population and mean that millions fewer teens have engaged in alcohol misuse since 1997.

The authors state that short of enforcement of the minimum drinking age, little has been done over the years to reduce underage drinking. They also suggest a number of interventions that should be implemented to reduce underage drinking. I respectfully disagree with their assertions and admonitions. I think it is critical that researchers carefully study these positive trends and try to determine what factors have actually contributed to substantial reductions in youth binge drinking. A better understanding of the reasons behind these positive findings will eventually lead to evidence-based interventions and public policy.

In the meantime, physicians caring for teens should be aware that a substantial minority of them (~25%–28%) have likely engaged in alcohol misuse in the preceding month. Screening, counseling, and appropriate referral are vital interventions that can be conducted in the office setting and will likely reduce future high-risk alcohol use among these patients. Perhaps as important, for alcohol users and nonusers alike, clinicians should strongly reinforce that the majority of American teens (71%–75%) chose not to use alcohol (or to use it minimally). Adolescents generally respond to pressure to engage in normative social behaviors; they just need to know what the normative behavior is.

James C. Turner, MD
Department of Student Health, the National Social Norms Institute, and Department of Clinical Medicine
University of Virginia
Charlottesville, VA 22908

REFERENCES

In Reply.—

In response to Dr Turner, underage drinking remains common and dangerous. Alcohol is a leading preventable cause of death among young adults and contributes to the 3 leading causes of death among adolescents.3 In addition, youth drinking and binge drinking contribute to a number of social problems, including unintended pregnancy, violence, sexually transmitted infections, altered brain development, and subsequent alcohol problems in adulthood.2,4 Although rates of underage drinking have declined slightly among male high school students during the past decade, it is hard to take solace from our study when almost half of all US high school students drank alcohol during the previous 30 days and approximately two thirds of those individuals binge drank, typically on multiple occasions.

Although binge drinking is an extremely dangerous pattern of alcohol consumption, our report and other studies have shown that any alcohol consumption among youth is riskier than no consumption. For example, we found that high school students who drank but did not binge drink were 2 to 4 times more likely than nondrinkers to ride with a drinking driver, be sexually active, smoke cigarettes, and get into a fight. In addition, because respondents in adult population-based surveys tend to underreport how much they drink, it is also likely that some students who reported drinking at less-than-binge levels may have been misclassified because of potential reporting bias.4,5 Furthermore, youth generally weigh less than adults and, therefore, are more likely to attain impairment-level blood alcohol concentrations when drinking fewer than 5 drinks. Finally, underage drinkers are far more likely than adult drinkers to suffer alcohol-related harms (eg, motor vehicle crashes) at blood alcohol concentrations below 0.08%.6

Dr Turner’s glass-half-full paradigm reflects an


doi:10.1542/peds.2007-0344
approach called social norms theory, which postulates that educating adolescents or college students about normative behavior (ie, emphasizing that many people do not drink) will “pressure” them to behave in less risky ways. Although this approach has been endorsed by some educators, there are at least 3 concerns. First, according to the Guide to Community Preventive Services, there is insufficient evidence to determine whether social norming is effective in reducing alcohol-related outcomes among youth (eg, drinking and driving). Second, social norming programs are aimed primarily at individuals and do not address important societal determinants of youth drinking behavior, such as adult alcohol misuse and the larger alcohol-policy environment. Third, programs that have a large focus on social norming may have the unintended consequence of minimizing the public’s appreciation for the problem of underage drinking, thus eroding support for other interventions that are known to be effective and that require a strong public commitment.

Social attitudes about the acceptability of underage drinking do need to change, but minimizing the problem will only increase our complacency about solving it. The time has come to match the scope and intensity of the underage drinking problem with robust, population-based intervention policies recommended by the Institute of Medicine, including increased alcohol excise taxes, adequate enforcement of minimum legal drinking age laws, and additional reductions in alcohol marketing to youth-oriented audiences. We also need to recognize the strong relationship between binge drinking by adults and binge drinking by youth and implement effective strategies (eg, limiting alcohol-outlet density) to prevent binge drinking in the general population. Absent such policy changes, the impact of future clinical and educational efforts to reduce underage and binge drinking and to change social attitudes regarding this behavior will be limited.

Jacqueline W. Miller, MD
Timothy S. Naimi, MD, MPH
Robert D. Brewer, MD, MSPH
Centers for Disease Control and Prevention
Atlanta, GA 30333

REFERENCES

Neonatal Blue-Light Phototherapy Could Increase the Risk of Dysplastic Nevus Development

To the Editor.—

The number of individuals with large numbers of common and atypical melanocytic nevi has been continuously increasing in several white populations. Subjects with dysplastic nevi are at a substantially increased risk for the development of both uveal and cutaneous malignant melanoma.

Numerous epidemiologic data have demonstrated that sunlight exposure is the major environmental factor involved in the development of melanoma, and it also exerts a considerable influence on the nevus count of an individual. Blue-light phototherapy has been widely and successfully used for the treatment of neonatal jaundice to reduce the plasma concentration of bilirubin and, thus, to prevent kernicterus. So far, little is known about the long-term impact of neonatal phototherapy on nevus development.

In 2002 and 2003, 747 schoolchildren aged 14 to 18 years were investigated to determine the prevalence of common and atypical melanocytic nevi. Data were recorded with regard to the neonatal history of the students, such as prematurity, neonatal jaundice and blue-light phototherapy; 44.6% of the children had received phototherapy for the treatment of neonatal jaundice. Our results revealed that the prevalence of common melanocytic nevi was quite similar in the treated and untreated children, but exposed subjects were likely to have multiple (>100) moles. Neonatal blue-light phototherapy was associated with a significantly higher prevalence of clinically atypical nevi ($\chi^2 = 4.08$; degrees of freedom = 1; $P = .0433$ [Statistica 7.1; StatSoft, Inc,
The prevalence of dysplastic nevi was 19.1% in the untreated group and 25.2% in the treated group. Blue-light phototherapy resulted in a relative risk of 1.32 for the development of dysplastic nevi (odds ratio: 1.43; 95% confidence interval: 1.010–2.026). Bauer et al have reported that phototherapy of neonatal jaundice using a blue-light lamp is not associated with an increased risk of development of melanocytic nevi in children aged 2 to 7 years. Common melanocytic nevi first appear in early childhood, whereas dysplastic nevi arise in late childhood, usually around puberty. Accordingly, we focused on this age group in our study.

Because having a clinically atypical nevi is the most important independent phenotypic risk factor for the development of malignant melanoma, our data highlight the need for the dermatologic screening of children with a history of neonatal phototherapy. Phototherapy with blue-light lamps is a standard and essential therapeutic modality in neonatal care; therefore, additional studies are necessary to establish its potential long-term adverse effects.

In Reply.—

Csoma et al evaluated 747 teenagers and correlated their findings of dysplastic melanocytic nevi with a past history of “neonatal blue-light phototherapy.” Because dysplastic nevi increase the risk for the development of malignant melanoma, these and other recent observations might be concerning. On the other hand, a large study found no relationship between exposure to phototherapy and the development of melanocytic nevi in children aged 2 to 7 years.

It is difficult to tell how the study population was selected. Was this a random or convenience population of 747 schoolchildren, or were they being treated by a dermatologist because of the presence of atypical nevi? Obtaining an accurate history of phototherapy exposure from parents and, in particular, the type of phototherapy used when 14 to 18 years have elapsed is certainly difficult. Are the authors confident that the exposure was to blue light and not broad-spectrum or tungsten-halogen phototherapy? In the United States in the 1980s, the vast majority of infants who received phototherapy were exposed to daylight or cool white lights, not blue or special blue lights. Perhaps the situation was different in Hungary.

If this was an unselected group of schoolchildren, the fact that 44.6% were said to have received phototherapy is remarkable. In a similar population in the United States, exposure to phototherapy would be expected to be <5%.

Exposure to UV light is considered to be an important risk factor for the development of malignant melanoma. Fluorescent bulbs do emit some UV-A light (315–400 nm), but most of the higher-energy (shorter-wavelength) part of that is blocked by the phosphor and glass envelope. If properly used with a Plexiglas shield, all radiation below ~340 nm would be removed.

We agree that ongoing studies are necessary to identify any potential long-term effects of phototherapy, but it is premature to suggest that children with a history of neonatal phototherapy should undergo “dermatologic screening.” A recommendation that calls for dermatologists worldwide to examine hundreds of thousands of children annually requires much stronger evidence.

ACKNOWLEDGMENT

This study was supported by the National Fund of the Hungarian Ministry of Health (ETT 500/2006 grant).

REFERENCES


doi:10.1542/peds.2007-0180

ACKNOWLEDGMENT

We thank Antony F. McDonagh, PhD, for his advice.
REFERENCES


doi:10.1542/peds.2007-0452

Kernicterus, the Daubert Decision, and Evidence-Based Medicine

To the Editor.—

The special article by Sartore and van Doren and the commentary by Brent draw attention to the Daubert decision that a court is required to decide whether medical expert testimony is scientifically valid. Notwithstanding the reservations expressed by Brent, the scientific community, generally, will applaud this decision. The authors also note that Daubert requires evidence from epidemiologic studies establishing a relative risk of >2.0 before the jury can conclude that there is a >50% likelihood that a putative cause produced the undesirable outcome. They then state that a Daubert motion could succeed in barring testimony by experts in cases of hyperbilirubinemia because there is “a paucity of medical literature to establish the relative risks of bilirubin-induced encephalopathy for neonates exposed to various levels or doses of bilirubin.”

Although quantitative literature on this point is admittedly sparse, the causal relationship between hyperbilirubinemia and kernicterus is generally not disputed. In our experience, however, most malpractice cases involving kernicterus do not revolve around the question of whether a specific bilirubin level produced brain damage or whether an intervention at a lower bilirubin level would have prevented the damage. More commonly, cases involve infants admitted to the hospital with bilirubin levels of ≥35 mg/dL who manifest the classical signs of acute bilirubin encephalopathy. In these cases, the causal role of hyperbilirubinemia is not in doubt and, in most such cases, it can be asserted with “reasonable medical probability” that, if this infant had been seen and treated earlier, the bilirubin would not have risen to dangerous levels and the kernicterus could have been prevented. Whether someone is to blame for this sequence of events, however, is often the more contentious issue, and specifically whether there was a breach in the standard of care.

We certainly want the court to assess the scientific validity of all expert testimony and advise the jury accordingly. Unfortunately, in many kernicterus malpractice cases, the jury is not asked to decide whether the particular bilirubin level was responsible for the brain damage; they are asked to decide if the physician’s actions complied with the prevailing “standard of care.” In legal parlance this is generally defined as what a reasonable physician would do with the same or similar patient under the same or similar circumstances. Because scientific evidence on this question is generally limited, the jury is often left to decide which expert(s) to believe.

M. Jeffrey Maisels, MB, BCh
Department of Pediatrics
William Beaumont Hospital
Royal Oak, MI 48073

Thomas B. Newman, MD, MPH
Departments of Epidemiology and Biostatistics and Pediatrics
University of California
San Francisco, CA 94143-0560

REFERENCES


doi:10.1542/peds.2007-0155

In Reply.—

Unlike the cases that Maisels and Newman describe, the cases that inspired our article involved critically ill newborns with multiple severe problems. Total serum bilirubin (TSB) levels peaked at slightly above 20 mg/dL but did not approach the level of 35 mg/dL. The physicians did not perform exchange transfusions, and the infants developed bilirubin-induced encephalopathy (BIE). These were severely ill infants who were still being treated in the NICU after several days of life. They were sedated, on ventilation, and receiving vasopressors and multiple antibiotics. Defense experts reported that the physicians did not breach standards of care in deciding not to perform exchange transfusions.

At trial the judge would have had a duty to instruct the jury on 2 issues relating to liability. The first issue is whether the physician met the standard of care in not performing exchange transfusion, and the second is whether an exchange transfusion would have prevented the BIE. A plaintiff must prevail on both issues to establish liability.

In a recent case, the Daubert line of cases supported a defense motion to exclude expert testimony for the plaintiff to the effect that an exchange transfusion performed when the TSB concentration was in the range of 16 to 18 mg/dL would probably have prevented the TSB level from exceeding 20 mg/dL and would have prevented BIE. The defense argued that the relative risk of BIE for exposure to different levels of TSB within a range of 16 to 22 mg/dL has not been determined by studies that meet Daubert standards (ie, the hypothesis for specific relative risks has not been subjected to testing, has not been peer reviewed and published, does not have a
known rate of error, and is not generally accepted in the scientific community). The plaintiff’s expert admitted this. Therefore, the defense argued, the expert could not conclude on the basis of relative risks that exposure to a TSB concentration >20 mg/dL was more likely the cause of BIE than exposure to a TSB level in the range of 16 to 20 mg/dL. The case was settled before the motion was decided.

Daubert and the cases that have come after it have the salutary effect of bringing evidence-based medicine to the court room. The purpose of our article was to introduce this concept to practicing physicians.

Both Extremes of Arterial Carbon Dioxide Pressure and the Magnitude of Fluctuations in Arterial Carbon Dioxide Pressure Are Associated With Severe Intraventricular Hemorrhage in Preterm Infants

To the Editor.—

I read the recent article by Fabres et al1 with great interest. The authors of the study observed that hypocapnia, extreme hypercapnia, and fluctuations of PaCO2 in premature infants with birth weights of 401 to 1250 g during the first 4 days of life are associated with severe intraventricular hemorrhage (IVH). The authors correctly pointed out that avoiding extremes of PaCO2 during the period when these infants are at highest risk for developing IVH may be prudent. I am pleased that their work nicely expands on the observations from my study that hypercapnia during the first 3 days of life is associated with severe IVH in very low birth weight infants.2

Disturbed cerebral blood flow and cerebral autoregulation are considered to play important roles in the development of IVH.3–5 With respect to the role that hypercapnia may play in the etiology of severe IVH, my study group was the first to show in premature infants during the first week of life that increasing PaCO2 is associated with progressively impaired cerebral autoregulation.6 Thus, because PaCO2 is an important regulator of cerebral blood flow,7 we suggest a plausible mechanism for the association of early hypercapnia and development for IVH.

I compliment Fabres et al1 because their work importantly adds to the burgeoning literature on the potential negative effects of extremes of Paco2 during the early neonatal period in extremely premature infants.

Jeffrey R. Kaiser, MD, MA
Division of Neonatology
Departments of Pediatrics and Obstetrics and Gynecology
College of Medicine
University of Arkansas for Medical Sciences
Little Rock, AR 72205

ACKNOWLEDGMENT
This research was supported by National Institute of Neurological Disorders and Stroke grant 1 K23 NS43185.

REFERENCES
2. Kaiser JR, Gauss CH, Pont MM, Williams DK. Hypercapnia during the first 3 days of life is associated with severe intraventricular hemorrhage in very low birth weight infants. J Perinatol. 2006;26:279–285

In Reply.—

We thank Dr Kaiser for his comments on our article.1 His comments focus on the association of hypercapnia with severe (grade III and IV) intraventricular hemorrhage (IVH), and his demonstration that increasing Paco2 is associated with impaired cerebral autoregulation in premature infants2 suggests a possible mechanism for the IVH resulting from extreme hypercapnia. However, there were no significant differences in Paco2 in infants with mild (grade I or II) IVH compared with infants with no IVH, which suggests that abnormal levels of Paco2 are more likely to be associated with extension of preexisting hemorrhage rather than initiation or development of IVH.1 Dr Kaiser’s comments also focus only on part of our findings. Our study also demonstrated an association of severe IVH with hypocapnia and fluctuations of Paco2.1 Previous studies have primarily demonstrated an association between periventricular leukomalacia and
hypocapnia,\textsuperscript{3,4} which may be caused by decreased cerebral blood flow and increased cerebral fractional oxygen extraction induced by the hypocapnia.\textsuperscript{5} As mentioned in our discussion, the mechanisms that underlie the association of severe IVH with hypocapnia and marked fluctuations in PaCO\textsubscript{2} need to be determined in future studies, although decreases in cerebral blood flow induced by hypocapnia and ischemia-reperfusion resulting from marked fluctuations in PaCO\textsubscript{2} may possibly be involved. Alternatively, hypocapnia, hypercapnia, and the fluctuations in PaCO\textsubscript{2} could be effects, rather than the causes, of severe IVH, by either fluctuations in respiratory drive or changes in carbon dioxide production caused by intracranial pathology.

Namasivayam Ambalavanan, MD
Departments of Pediatrics
University of Alabama at Birmingham
Birmingham, AL 35249

REFERENCES
1. Fabres J, Carlo WA, Phillips V, Howard G, Ambalavanan N. Both extremes of arterial carbon dioxide pressure and the magnitude of fluctuations in arterial carbon dioxide pressure are associated with severe intraventricular hemorrhage in preterm infants. 
Pediatrics. 2007;119:299–305
2. Kaiser JR, Gauss CH, Williams DK. The effects of hypercapnia on cerebral autoregulation in ventilated very low birth weight infants. 
Pediatri Res. 2005;58:931–935
Pediatrics. 1996;98:918–924
Pediatri Res. 2005;58:579–585

Deterrent to Healthy Lifestyles in Our Communities

To the Editor.—

We certainly agree with the recent letter to the editor of Pediatrics\textsuperscript{1} that emphasized that the current epidemic of obesity requires prevention work on several fronts including the identification of barriers to healthy lifestyles in our communities. In addition, we also believe that it is important to identify best practices of obesity prevention irrespective of whether they are active or passive. However, the author raised a very important perspective: “Our reliance on the automobile and our propensity to design our urban environments for cars rather than for people may well be an additional root cause of the obesity epidemic.”\textsuperscript{1} Transport is an important determinant of public health. It affects physical activity levels, crashes and injuries, social and economic opportunities, and general well-being.\textsuperscript{2} Extensive literature on the various health benefits of physical activity has necessitated the understanding of determinants, risks, and barriers to physical activity.

It has been highlighted that one of the biggest epidemics facing the world in the 21st century is the growing obesity in children. This epidemic has been attributed largely to sedentary lifestyles. Therefore, it is critical to identify the barriers and the potentially effective strategies for surmounting the problems that hinder walking and cycling.\textsuperscript{3}

In 1969, approximately half of all schoolchildren in the United States walked or bicycled to or from school. Today, fewer than 15% of children and adolescents use these active modes of transportation.\textsuperscript{4} In recent years, new public health challenges have brought the issue of the decline of young people’s physical activities to the forefront. One recent study in the United States explored the question of why children do not walk to school more often. Parents reported multiple barriers that inhibit walking and biking to school as follows: long distances (55%), traffic danger (40%), weather (24%), crime (18%), and school policy (7%). Similarly, in the United Kingdom a recent study showed that 85% of parents were worried about traffic dangers on their children’s journey to school.\textsuperscript{5}

Parents discourage their children from walking or cycling to school because they worry about the dangers from traffic.\textsuperscript{6} As a result, less than half primary school–aged children are now walking to school in the United Kingdom.\textsuperscript{7}

Many neighborhoods lack sidewalks, bike lanes, and safe paths or have dangerously fast traffic. Barriers to routine physical activity have sprung up nearly everywhere in the lives of children. At the same time, children have an ever-increasing set of sedentary indoor entertainment to keep them occupied. Although public health policies encourage more children to walk and cycle to school, this can only happen when parents’ rightful worries about their children’s safety are addressed adequately.

A recent study shows that in the 5- to 9-year-old age group, sport utility vehicles (SUVs) were 4 times as likely to be associated with fatal pedestrian injury (odds ratio: 4.2; 95% confidence interval: 1.9–9.5). It was concluded that vehicle body-type characteristics play an important role in fatal pediatric pedestrian injuries.\textsuperscript{8} Reducing SUVs’ frontal design seems then to be a worthwhile public health endeavor.

Road danger is a real disincentive to active transport, and reducing traffic-injury risks for child pedestrians and cyclists should be an important part of any strategy to encourage walking and cycling, which then will assist in
increasing active lifestyles in our communities and subsequently reducing childhood obesity. A recent study in the United Kingdom suggested that more needs to be done in this respect.9 If children avoid walking and bicycling because of parental fears (justified or not) of them being injured by motor vehicles, then the disease burden includes inactivity, and the total burden of disease will be much larger than the World Health Organization predicts for 2020.10

In 2001, an estimated 685 000 children under the age of 15 were killed by unintentional injuries. Approximately 20% of all global deaths from unintentional injuries occur in this age group, which makes unintentional injury among the 10 leading cause of death in our children.11

Are our driveways even safe? Children playing in driveways are at risk of fatal and serious injury caused by reversing vehicles.12 Driveway injuries are an underrecognized, often severe form of motor vehicle–pedestrian crashes. Injuries within residential driveways occur most often to children younger than 5 years and carry a significant risk of injury and mortality. Recent research shows that with the increasing popularity of SUVs, minivans, and large-model pick-up trucks, children involved in driveway back-over fatal injuries are increasing in the United States and Australia.13–15

Similarly a recent study in Brazil showed that many young children playing in driveways are often killed by SUVs.16 Parks and playgrounds are the most obvious locales for creating active environments for kids, yet in developing countries there is a dearth of safe outdoor facilities. One recent study showed that increasing the number of public parks correlates with a decreased rate of motor vehicle–related mortality in young children.17

The vast majority of unintentional injuries among children occur in low- and middle-income countries, and children in African, southeast Asian, and Western Pacific regions account for 80% of all deaths from unintentional injuries.18 Coupled with the lack of resources available for unintentional injury prevention and control in developing countries, this increase in motor vehicle danger portends a holocaust of preventable child deaths.

Systematic reviews have suggested that traffic-calming schemes do reduce the number of accidents by ~15%,19 and a recent Cochrane review20 of area-wide traffic-calming in towns and cities suggested that it may be a promising intervention for reducing the number of road-traffic injuries and deaths, with a pooled rate ratio of 0.89 (95% confidence interval: 0.80–1.00).

Much promising current research shows that the level of pedestrian protection provided by motor vehicles can be significantly improved with practicable design changes. Road deaths can be greatly reduced by improving vehicle design; however, the automobile industry is resisting this in favor of style and speed.21

Therefore, policy-makers need to do work that would require automakers to design more pedestrian-friendly vehicles. Pediatricians should consider any work in vehicle design that aims to reduce the injury severity to child pedestrians to be vitally important.

**REFERENCES**


In Reply.—

I am gratified that the group from British Columbia has expanded on my comment and pointed out the distressing hazards to child safety posed by both vehicles and the roadways they occupy, the design of which seldom shows much thought for pedestrians or cyclists. Long distance and parents' fear of crime are other notable barriers to walking or biking to school that they cite. Long distance is inherent in urban design subservient to automobiles, where the fabric must be stretched just to make room for cars, their storage, and the possibility that they can move at all.

Parents' fear of crime suggests another overarching issue precipitated by our automobile cities: the loss of social capital. Social capital can be defined as mutual trust, psychological sense of community, neighborhood cohesion, and community competence. Jane Jacobs, an early influential scholar of American urban design, argued that the active pedestrian street life of older neighborhoods with shops, small streets, and mixed uses promoted safety. She coined the term “eyes on the street” and noted the sense of community in these neighborhoods. The concern of inhabitants for their fellows was in stark contrast to the high-rise projects and streets empty except for cars, which makes robbery and personal attack easier.2

Urban sprawl adds to the problem by segregating neighborhoods by income, taking good schools, jobs, and other resources to the periphery and leaving behind poverty and pollution in the inner city. Commentary on pollution generated by automobiles, profligate water and energy use, and global warming, all exacerbated by urban sprawl, is beyond the scope of this letter. However, one further issue might be mentioned as an example. Asthma remains a burdensome pediatric problem, and that burden is borne disproportionately by black children. A study from Hartford, Connecticut, an urban minority community, used the health-field concept and studied factors that determine the health of communities: environment, human biology (genetics), personal behavior, and health care organization. The study concluded that “[i]mproved personal behaviors and medical care will have a limited sustained impact on childhood asthma until basic environmental issues are modified.”4

Approaches to modifying environmental issues can be small or big. A neighborhood in Portland, Oregon, created a gathering place at a formerly busy intersection by painting a giant sunflower across the intersection, installing a fountain, an art wall, and an information kiosk. Walking and biking increased, as did measures of social capital and a neighborhood sense of well-being.5

The big approach is illustrated by the urban-design principles advocated by the New Urbanism, including walkability, mixed use and diversity, mixed housing, traditional neighborhood structures, increased density, and smart pedestrian-friendly transportation. Big or little approaches to problems of the built environment could go a long way in mitigating a number of health issues, including obesity.

James Fisk, MD
Department of Pediatrics
University of New Mexico
Albuquerque, NM 87131

REFERENCES

doi:10.1542/peds.2007-0635

Autologous Cord Blood Transplantation in a Child With Acute Lymphoblastic Leukemia and Central Nervous System Relapse

To the Editor.—

We read with great interest “First Report of Autologous Cord Blood Transplantation in the Treatment of a Child With Leukemia” by Hayani et al.1 The authors reported on the first successful autologous cord blood transplantation in a child with acute lymphoblastic leukemia of the B-precursor subtype with isolated central nervous system (CNS) relapse within 10 months from diagnosis. However, CNS and total-body irradiation (TBI) combined with high-dose conditioning and autologous stem cell transplantation can be considered an overtreatment in a 3-year-old child with acute lymphoblastic leukemia.
and CNS relapse occurring within 18 months after initial diagnosis, because event-free survival is not superior in these patients compared with patients who are treated less intensively.2,3 It is well known that umbilical cord blood (UCB) is associated with higher rates of engraftment failure compared with peripheral blood stem cells, which could have been used as well assuming that both sources are equally effective. Because TBI is myeloablative and UCB neither provides the opportunity to boost delayed engraftment nor to harvest stem cells for a second time in case of graft failure, at least autologous peripheral blood stem cells should have been harvested as a backup in case of nonengraftment of the UCB. In addition, the long-term adverse effects of TBI such as infertility, growth retardation, and an increased risk of secondary neoplasms should have been kept in mind. We consider it a severe conflict of interest that one of the authors employed at the commercial UCB bank where the autologous cord blood was stored advertises this case report on the company’s homepage as proof of principle for treating relapsed leukemia. Furthermore, this case report is now being advertised uncritically in many other private UCB banks worldwide.

In Reply.—
In general, central nervous system relapse is not a clear and absolute indication for hematopoietic stem cell transplantation in children with acute lymphoblastic leukemia. However, early central nervous system relapse is associated with poor outcome when treated with conventional chemotherapy and radiation therapy. Although there have not been any randomized studies to establish an advantage of stem cell transplantation, such transplantation seems to provide better survival rates compared with conventional therapy (77% vs 46%).1,2 In our patient, the very early onset of relapse (only 10 months) and the fact that she relapsed while receiving a more intensive chemotherapy regimen for high-risk leukemia put her at a high risk for treatment failure if she had been treated with conventional therapy. These factors led us to favor hematopoietic stem cell transplantation over conventional chemotherapy and radiotherapy.

Given the fact that the patient was still relatively young at the time of transplant and the viability of cord blood was excellent, we felt that the risk of graft failure was small, and collection of backup stem cells was not necessary. Autologous stem cells would have also had the possibility of being contaminated with residual leukemia cells.

We played no role in the family’s decision to collect and save our patient’s cord blood, and our report does not advocate private cord blood collection. One case report does not establish the efficacy or safety of autologous cord blood transplantation. It would be considered inappropriate for anyone to use this report as an argument in favor of private cord blood collection. Our main objective in writing the report was to discuss feasibility, safety, and uncertainties of autologous cord blood transplantation. We do believe that for children with high-risk relapsed leukemia in whom cord blood has been collected, the option of autologous cord blood transplantation should be considered and its risks and benefits should be weighed against conventional therapy and allogeneic stem cell transplantation.

REFERENCES
Improved Outcomes of Extremely Low Birth Weight Infants

To the Editor.—

It was with much interest that we read the article by Wilson-Costello et al that reported the improved neurodevelopmental outcomes for extremely low birth weight (ELBW) infants born in 2000–2002. This cohort included nearly 1500 children overall, with 233 children born in 2000–2002. More than 92% of the surviving children born in 2000–2002 were tested at follow-up. Results of this large study revealed a relatively stable survival rate from 1990–1999 compared with 2000–2002 (68% vs 71%, respectively) and a significant decrease in neurodevelopmental impairment from 35% in the 1990s to 23% in 2000–2002. The rate of cerebral palsy decreased from 13% in the 1990s to a 5% rate in 2000–2002. These findings were attributed to a variety of factors including increased antenatal steroid use and cesarean-section delivery as well as decreased sepsis, severe cranial ultrasound abnormalities, and postnatal steroid use.

Although our sample size is smaller and spans the years 1998–2000, we too have found encouraging neurodevelopmental outcomes for ELBW infants at 6 years of age. We report a survival rate of 80% to 85% for the years 1998–2000. Our ELBW cohort had average function across a wide range of standardized cognitive, neuropsychological, and behavioral tests, with no subjects falling out of the reference range for any test administered. Standardized tests of general intelligence, academic achievement, executive function, attention, language, memory, motor function, and behavior were administered. Infants born at 23 to 25 weeks’ gestation had no significant differences in outcome compared with those born at 26 to 28 weeks’ gestation. It is interesting to note that our cohort had a 75% incidence of sepsis during their neonatal course, and 100% received postnatal dexamethasone therapy for chronic lung disease. None of our subjects had greater than a grade II intraventricular hemorrhage. Our data support the findings of improved neurodevelopmental outcome for ELBW infants, although we cannot attribute these findings to lack of dexamethasone use postnatally or to a decreased sepsis incidence. Our data suggest that in the absence of significant intracerebral complications, at least average early school age outcomes are likely regardless of gestational age, birth weight, and use of postnatal steroids. In addition, these infants can be expected to have average neurodevelopmental outcome regardless of common medical complications during the newborn period and that birth before 26 weeks’ gestational age does not portend intellectual, neuropsychological, or behavioral deficit.

Margot Ahronovich, MD
Ida Sue Baron, PhD
Fern Litman, MD
Department of Neonatology
Inova Fairfax Hospital for Children
Falls Church, VA 22042

REFERENCES

doi:10.1542/peds.2007-0569
ERRATA

Dickerman JD. The Late Effects of Childhood Cancer Therapy. PEDIATRICS 2007;119:554–568.

Several errors occurred in the article by Dickerman titled “The Late Effects of Childhood Cancer Therapy” published in the March 2007 issue of Pediatrics (doi:10.1542/peds.2006-2826). In The Problem section on page 555, line 36, the author wrote: “legally blind loss of an eye.” It should read: “legally blind or loss of an eye.” In the Thyroid section on page 558, line 27, the author wrote: “7.2 to 20.4 times that subjects without Hodgkin’s lymphoma.” It should read: “7.2 to 20.4 times that seen in subjects without Hodgkin’s lymphoma.” In Table 2 on page 563, under the heading “Abdominal/pelvic radiation,” line 8, the word “Cancer” should be indented. On line 9 of the same section, the word “Bladder” should not be indented.

doi:10.1542/peds.2007-0753
Dr. Ralph Feigin was born in New York City on April 3, 1938. He received his A.B. degree from Columbia College in New York City in 1958 and his M.D. degree from Boston University School of Medicine in 1962. He completed a Pediatric Internship at the Boston City Hospital from 1962–63 and served as a Pediatric Resident at the Boston City Hospital from 1963–64 and at the Massachusetts General Hospital from 1964–65. He subsequently completed a Research Assignment with the United States Army Research Institute of Infectious Diseases in Frederick, Maryland from 1965–67. From 1967–68 he served as Chief Resident of the Children’s Service at the Massachusetts General Hospital. Dr. Feigin was a member of the faculty of Washington University School of Medicine in St. Louis, Missouri from 1968 to 1977 when he was serving as Professor of Pediatrics, Director of the Division of Infectious Diseases in the Department of Pediatrics and Director of the Bacteriology and Serology Laboratories at the St. Louis Children’s Hospital.

In July of 1977, Dr. Feigin was appointed and currently serves as the J.S. Abercrombie Professor of Pediatrics and Chairman of the Department of Pediatrics at Baylor College of Medicine and Physician-in-Chief of the Texas Children’s Hospital. In addition, he is Physician-in-Chief, Pediatric Services, Ben Taub General Hospital (Harris County Hospital District), and Chief of the Pediatric Service, The Methodist Hospital, Houston, Texas. In 1992 he was appointed Senior Vice President of Baylor College of Medicine and in 1994 he was appointed Dean of Medical Education for the School, positions he held until his appointment as President and Chief Executive Officer of the Baylor College of Medicine in January of 1996. He served as President and Chief Executive Officer until March of 2003.

Dr. Feigin was elected to membership in Alpha Omega Alpha, the National Honor Medical Society, in 1961. He was the recipient of an USPHS Research Career Development Award from the National Institute of Allergy and Infectious Diseases in 1970. He received the Senior Class Award to the Outstanding Teacher of the Year from Washington University School of Medicine in 1975, was recognized as an Alumni Teaching Scholar at Washington University School of Medicine in 1975, and was the recipient of the Founders Day Award from Washington University School of Medicine in 1977. He has received the Senior Class Outstanding Teacher Award from Baylor College of Medicine in 1979, 1980, 1981, 1982, 1983 (both the May and November graduating classes), 1984, 1985 and 1986. In 1984 he was elected to the Outstanding Faculty Hall of Fame of Baylor College of Medicine. In 1989 Dr. Feigin received the Distinguished Alumnus Award from Boston University School of Medicine. In 1995 Dr. Feigin received the Joseph W. St. Gme, Jr., Leadership Award and, also in 1995, he was elected to membership in the National Academy of Sciences Institute of Medicine. He received the Distinguished Physician Award from the Pediatric Infectious Disease Society in 1996 and was the recipient of the 1997 American Academy of Pediatrics Medical Education Lifetime Achievement Award. In 1998 Dr. Feigin received an honorary degree, Doctor of Humane Letters, from Boston University at the 150th Anniversary Celebration of the Boston University School of Medicine. Dr. Feigin is the recipient of the 2007 John P. McGovern Compleat Physician Award.

Dr. Feigin is a member of the Society for Pediatric Research, of which he served as President from 1982 to 1983, the American Academy of Pediatrics, the Infectious Diseases Society of America, the American Pediatric Society, of which he served as President from 1997 to 1998, and the Association of Medical School Pediatric Department Chairmen, of which he served as President from 1991 to 1993. He is a member of numerous other national, state and local organizations, including have served as a member of the Board of Governors and Finance Committee of the National Institutes of Health Warren Grant Magnuson Clinical Center from 2000 to 2005.

Importance of the Clinical Recognition of Loeys-Dietz Syndrome in the Neonatal Period

Anji T. Yetman, MDa, Rebecca S. Beroukhim, MDb, Dunbar D. Ivy, MDb, David Manchester, MDb

aDepartment of Pediatrics, Divisions of Cardiology and Genetics, Children’s Hospital, University of Colorado Health Sciences Center, Denver, Colorado; bDepartment of Pediatrics, Division of Cardiology, Primary Children’s Medical Center, University of Utah, Salt Lake City, Utah

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

ABSTRACT

We describe 5 patients who presented with musculoskeletal abnormalities in the neonatal period. All patients were initially suspected to have Larsen syndrome or Beals syndrome but were subsequently diagnosed with a TGFBR2 mutation diagnostic of Loeys-Dietz syndrome. Patients had progressive aortic enlargement, which necessitated surgical intervention for 3 patients and resulted in the death of 1 patient. Delay in diagnosis of Loeys-Dietz syndrome may be associated with adverse prognosis.

IN 2005, LOEYS et al1 reported on a newly identified genetic syndrome, Loeys-Dietz syndrome. The syndrome shares many clinical characteristics with Marfan syndrome, including the propensity for early aortic dilation and dissection.1 The cellular mechanisms that underlie the aortic pathology in Loeys-Dietz syndrome are similar to Marfan syndrome, including altered activity of the cytokine family transforming growth factor β (TGF-β).2 Unlike patients with Marfan syndrome, those with Loeys-Dietz syndrome do not have an alteration in the FBN-1 gene encoding the fibrillin protein but, rather, a mutation in the serine-threonine kinase domain of 1 of 2 TGF-β receptors, TGFBR1 or TGFBR2.2 To date, reports from 2 series (totaling 52 cases) of patients with Loeys-Dietz syndrome have been published, with the majority of patients having reached adulthood.1,2 Given the very recent recognition of the syndrome, the natural history of the clinical features and the timing of presentation remain unknown.

Here we describe 5 patients with genetically confirmed Loeys-Dietz syndrome who presented for medical attention in the neonatal period with a clinical picture of joint contractures and joint hypermobility. Despite the fact that all the patients had progressive cardiac pathology during childhood, most had inadequate cardiac follow-up, because all had been misdiagnosed.

CASE REPORTS

Patient 1 was a term infant who was referred to a tertiary pediatric care center on day 1 of life for additional evaluation of a genetic syndrome in the setting of a murmur, diffuse hypotonia, macrocrania, and upper- and lower-limb abnormalities (Table 1). Weight was in the 3rd percentile, height was in the 10th percentile, and head circumference was >97th percentile. The patient was evaluated by the department of genetics, and a differential diagnosis including Larsen syndrome, Beals syndrome, and other “arthrogryposis complexes” was considered in light of the multiple joint dislocations and contractures. Family history was unremarkable. Initial investigations included normal chromosomal study results, skeletal imaging (Table 1), computed tomography (CT) of the head that demonstrated no intracranial abnormalities or abnormalities of the skull, and an echocardiogram that revealed a structurally normal heart, dilated pulmonary artery, dilated aortic root (Table 2), and a small patent ductus arteriosus (PDA) without aneurysm. Annual cardiology follow-up was arranged. Surgical interventions during infancy included repair of talipes equinovarus and metatarsus adductus foot deformities, umbilical hernia repair, and repair of bilateral exotropia. At 1 year of age, the patient was noted to...
have had significant progression in aortic dilation. She was placed on β-blocker therapy with no subsequent change in the rate of aortic dilation. She was subsequently referred to our Marfan subspecialty clinic, at which time she was started on an angiotensin-converting enzyme (ACE) inhibitor and followed every 3 months with echocardiography. Aortic dilation stabilized with no additional increase in percent-predicted aortic diameter. After the recent report in *Nature Genetics,* DNA sequencing was performed and confirmed the presence of a previously reported mutation in TGFBR2 (Table 3). At 2 years old, solely as a result of the genetic diagnosis, MRI of the cervical spine was obtained, which demonstrated spinal canal stenosis caused by cervical vertebral subluxation (Table 1). The patient was followed by neurosurgery and noted to have progressive spinal cord compression necessitating surgical intervention at 30 months old.

Patient 2 was a term infant who was referred to a tertiary pediatric center on day 1 of life for additional evaluation of a genetic syndrome in the setting of a murmur, diffuse hypotonia, macrocrania, and musculoskeletal abnormalities including bilateral knee dislocations and hip dislocations (Table 1). Weight was in the 97th percentile, height was in the 97th percentile, and head circumference was at >99th percentile. Because of the lower-extremity abnormalities and a finding of “abnormal facies” with hypertelorism, the patient was suspected to have Larsen syndrome. Family history was unremarkable. Investigations included normal chromosomal study results, skeletal imaging (Table 1), CT of the head that demonstrated no intracranial abnormalities, and an echocardiogram that demonstrated a large ductal aneurysm with thrombus, as well as pulmonary artery and aortic root enlargement (Table 2). Surgical interventions during the neonatal period included PDA ligation. Because of the ductal aneurysm and the other clinical features (Table 1), genetic studies for Loeys-Dietz syndrome were performed. DNA sequencing demonstrated a missense mutation in the kinase domain of TGFBR2 (Table 3). At the age of 10 years, the patient was noted to have progressive aortic root enlargement (Table 2) and was started on an ACE inhibitor. At 6 months of age the aortic root diameter was stable. The patient is being followed by neurosurgery for cervical vertebral abnormalities (Table 1).

Patient 3 was a term infant who was referred to a tertiary pediatric center on day 1 of life for apnea, diffuse hypotonia, and musculoskeletal abnormalities (Table 1). All growth parameters were within normal limits. Head circumference was in the 97th percentile. Because of the lower-extremity abnormalities and facial appearance, the patient was suspected to have Larsen syndrome. Family history was unremarkable. Investigations included normal chromosomal study results and skeletal imaging (Table 1). Surgical interventions during childhood included serial orthopedic interventions for bilateral talipes equinovarus, inguinal hernia repair, and repair of bilateral exotropia. The patient presented acutely for medical attention at 6 years of age with arm numbness and was noted to have spinal cord compression secondary to C2–C3 subluxation. He went on to require spinal fusion. The patient re-presented at 9 years of age with jaw pain and was found to have an aortic dissection in the setting of a markedly dilated aortic root. The dissection extended into the vertebral and carotid arteries and, despite successful surgical intervention, the patient was left with a diffuse hypoxic-ischemic cerebral insult that resulted in death. Because of the aortic dissection at a young age and the presence of a bifid uvula noted on intubation, the patient underwent genetic testing at the time of surgical intervention. DNA sequencing demonstrated a missense mutation in the kinase domain of TGFBR2 (Table 3).

Patient 4 was a term infant who was noted to have marked hypotonia and several musculoskeletal abnormalities (Table 1) that prompted an outpatient orthopedic evaluation. Growth parameters were within normal limits. The patient was diagnosed with arthrogryposis, and the question of Larsen syndrome was raised. Family history was unremarkable. Investigations included normal chromosomal study results and skeletal imaging (Table 1). Surgical interventions during childhood included correction of lower-extremity abnormalities, repair of bilateral strabismus, bilateral hernia repair, and repair of an acquired Morgagni hernia. Clinical course

---

**TABLE 1 Clinical Manifestations**

<table>
<thead>
<tr>
<th>Manifestation</th>
<th>Patient No.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Neck</strong></td>
<td></td>
</tr>
<tr>
<td>Age at diagnosis of C3 hypoplasia with</td>
<td>2y</td>
</tr>
<tr>
<td>vertebral subluxation ± cord compression</td>
<td></td>
</tr>
<tr>
<td><strong>Musculoskeletal: upper</strong></td>
<td></td>
</tr>
<tr>
<td>Radial-ulnar dislocation</td>
<td>+</td>
</tr>
<tr>
<td>Metacarpophalangeal dislocations</td>
<td>+</td>
</tr>
<tr>
<td>Camptodactyly</td>
<td>+</td>
</tr>
<tr>
<td><strong>Musculoskeletal: lower</strong></td>
<td></td>
</tr>
<tr>
<td>Talipes equinovarus</td>
<td>+</td>
</tr>
<tr>
<td>Metatarsus adductus</td>
<td>+</td>
</tr>
<tr>
<td>Knee dislocations</td>
<td>+</td>
</tr>
<tr>
<td>Hip dislocation</td>
<td>+</td>
</tr>
<tr>
<td>Calcaneal-cuboid dislocation</td>
<td>+</td>
</tr>
<tr>
<td>Herniae</td>
<td>+</td>
</tr>
<tr>
<td>Acquired diaphragmatic</td>
<td>+</td>
</tr>
<tr>
<td>Inguinal</td>
<td>+</td>
</tr>
<tr>
<td>Umbilical</td>
<td>+</td>
</tr>
</tbody>
</table>
was marked by persistent panligamentous laxity that required lower-body bracing to allow ambulation. The patient was noted to have a murmur during the first month of life and was noted to have a small restrictive PDA, a dilated pulmonary artery, and a dilated aortic root. The patient was subsequently lost to cardiology follow-up until re-presenting at 2 years of age with respiratory symptoms. Echocardiography at that time demonstrated a markedly dilated aortic root, which measured 51 mm, with mild aortic insufficiency (Table 2). The patient underwent ductal ligation and a valve-sparing aortic root-replacement procedure and was placed on an ACE inhibitor with a subsequently stable cardiac course. At 6 years of age, the patient began complaining of lower-extremity weakness. Cervical spine abnormalities were noted at that time (Table 1). Because of the cervical spine abnormality, genetic studies with DNA sequencing were performed and confirmed a TGFBR2 mutation (Table 3).

Patient 5 was a term infant who was referred to a pediatric center on day 1 of life for diffuse hypotonia, macrocrania, and multiple musculoskeletal abnormalities. Weight was at <3rd percentile, height was in the 5th percentile, and head circumference was at >98th percentile. The patient was suspected to have Beals syndrome. Family history was unremarkable. Investigations included normal chromosomal study results, fibrillin study results, skeletal imaging (Table 1), and CT of the head (which demonstrated no significant intracranial abnormalities or any abnormality of the skull with the exception of extremely shallow ocular orbits). At 2 months of age the patient presented with acute loss of consciousness. CT of the head and neck at that time demonstrated no abnormalities with the exception of an abnormal, hypoplastic C3 vertebra. The patient was followed closely in this regard and noted to have complete C2–C3 subluxation with spinal cord compression, which necessitated stabilization and subsequent fusion. At the time of orthopedic intervention, the patient developed supraventricular tachycardia, and a cardiology consult was performed. Subsequent echocardiography demonstrated a dilated aortic root. Because of rapidly progressive dilation, the patient underwent a valve-sparing aortic root-replacement procedure at 3 years of age with good results. Because of the cervical abnormalities and the aggressive aortic disease, the question of Loeys-Dietz syndrome was raised. DNA sequencing confirmed the diagnosis, with identification of a previously reported splice-site mutation in TGFBR2 (Table 3).

**DISCUSSION**

Loeys-Dietz syndrome is a newly recognized genetic condition. As we have demonstrated, patients with this syndrome may present in the neonatal period with a number of clinical features including PDA with or without ductal aneurysm, macrocrania without craniosynos-tosis, and joint contractures and/or marked ligamentous laxity.

In their original patient description, Larsen et al noted a combination of multiple joint contractures and dislocations and abnormal facial appearance including hypertelorism. However, several reports followed that noted the presence of cardiac abnormalities and cervical spine abnormalities in association with joint contractures and labeled this constitution of findings “Larsen syndrome.” The clinical overlap between these patients and those with genetically proven Loeys-Dietz syndrome suggests that Loeys-Dietz syndrome may be more common than initially realized. Misdiagnosis may not be uncommon. Recent discovery of the genetic abnormal-

---

**TABLE 2** Cardiac Manifestations

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Age at Diagnosis of Aortic Dilation</th>
<th>Current Age</th>
<th>% Aortic Roota</th>
<th>Arch Vessels</th>
<th>Other</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2 d</td>
<td>30 mo</td>
<td>147</td>
<td>Tortuous</td>
<td>PDA</td>
<td>Stable on medication</td>
</tr>
<tr>
<td>2</td>
<td>2 d</td>
<td>6 mo</td>
<td>140</td>
<td>Tortuous</td>
<td>PDA</td>
<td>PDA ligation; stable on medication</td>
</tr>
<tr>
<td>3</td>
<td>9 y</td>
<td>9 y</td>
<td>200</td>
<td>Tortuous</td>
<td>Aortic dissection</td>
<td>AVR with aortic replacement; death</td>
</tr>
<tr>
<td>4</td>
<td>1 mo</td>
<td>9 y</td>
<td>170</td>
<td>Tortuous</td>
<td>PDA</td>
<td>Valve-sparing surgery</td>
</tr>
<tr>
<td>5</td>
<td>6 mo</td>
<td>6 y</td>
<td>178</td>
<td>Tortuous</td>
<td>SVT</td>
<td>Valve-sparing surgery</td>
</tr>
</tbody>
</table>

SVT indicates supraventricular tachycardia; AVR, aortic valve replacement.

a Aortic root diameter as a percent of predicted on the basis of body surface area—established normal values. Value was obtained at last clinic follow-up before surgical intervention.

**TABLE 3** Genetic Test Results

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Mutation</th>
<th>Nucleotide</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TGFBR2 Exon 7</td>
<td>c. 1583G→A</td>
<td>Arg528His</td>
</tr>
<tr>
<td>2</td>
<td>TGFBR2 Exon 7</td>
<td>c. 1570G→A</td>
<td>Asp524Asn</td>
</tr>
<tr>
<td>3</td>
<td>TGFBR2 Exon 5</td>
<td>c. 1318G→A</td>
<td>Glu440lys</td>
</tr>
<tr>
<td>4</td>
<td>TGFBR2 Exon 4</td>
<td>c. 865–873delACAGAGAG</td>
<td>Thr289Lys291del</td>
</tr>
<tr>
<td>5</td>
<td>TGFBR2 IVSS–1G→A (splice-site mutation)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ities in both Larsen and Loeys-Dietz syndromes should help clarify this situation. Confirmatory genetic testing is imperative if either diagnosis is considered, because clinical follow-up will differ.

Beals syndrome, otherwise known as congenital contractual arachnodactyly, occurs as a result of a mutation in the fibrillin 2 gene. It shares some common features with Loeys-Dietz syndrome but is supposed to be devoid of cardiac or ocular involvement.9 Again, cases have been reported that involve aortic pathology, which raises questions regarding the accuracy of the diagnosis in these previous publications.9

Ductal aneurysms have been reported to occur in association with diffuse aortic enlargement and marfanoid features in both children10 and adults,11 which suggests that Loeys-Dietz syndrome should be part of the differential diagnosis for these patients as well.

In our cohort of patients, cervical spine abnormalities were universally present, and significant, which is a finding that was not previously reported. The lack of cervicospinal abnormalities in previously reported patients may represent clinical variability or absence of diagnosis in asymptomatic patients. We recommend imaging of the cervical spine in all patients who are suspected of having a TGFBR1 or TGFBR2 mutation.

Although the majority of previously reported patients with Loeys-Dietz syndrome have had diffuse aneurysmal formation, our patients, despite thorough magnetic resonance angiography evaluation, had aneurysmal involvement confined to the aortic root. There was a corkscrew appearance of the head and neck vessels in all patients, but there were no additional aneurysms of the arterial tree. The young age of our patient cohort may help explain this finding and suggests that patients need to be followed for developmental alterations in the vascular system over time.

CONCLUSIONS
Children who present with diffuse hypotonia, joint laxity, and/or joint contractures and in whom a diagnosis of arthrogryposis, Larsen syndrome, or Beal syndrome is being entertained should undergo thorough genetic evaluation to rule out a TGFBR1 or TGFBR2 mutation. Frequent cardiovascular imaging and imaging of the cervical spine are required for such patients.

REFERENCES
Diagnosis of Common Variable Immunodeficiency in a Patient With Primary Ciliary Dyskinesia

Edward W. Skorpinski, MD a, Shiang-Ju Kung, MD b, Ejaz Yousef, MD c,d, Stephen J. McGeady, MD a, d

aBerkshire Allergy & Asthma Center, Wyomissing, Pennsylvania; bAllergy Section, Division of Immunologic and Infectious Diseases, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania; cDivision of Allergy and Clinical Immunology, Alfred I. duPont Hospital for Children, Nemours Children’s Clinic, Wilmington, Delaware; dDivision of Allergy and Clinical Immunology, Thomas Jefferson University, Philadelphia, Pennsylvania

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

ABSTRACT

In this case report we describe the first account in the literature of a patient with primary ciliary dyskinesia and common variable immunodeficiency. A 17-year-old boy with previously diagnosed Kartagener syndrome and stable lung disease developed a deteriorating clinical course that prompted the search for a secondary diagnosis. Although both of these rare conditions can result in similar lung pathology, they require different management strategies, which illustrates the need to consider associated diagnoses in complicated clinical situations.

First described in 1933, Kartagener syndrome (KS) consists of the clinical triad of situs inversus, bronchiectasis, and recurrent sinusitis. In the 1970s, with the use of electron microscopy, the physiologic defect of KS was determined to be nonfunctioning cilia caused by absence of inner and outer dynein arms on the ciliary structure. Since that time, the term “primary ciliary dyskinesia” (PCD) has been used to describe the larger group of patients with recurrent infections and improperly functioning cilia, of which 50% have situs inversus. PCD is felt to be an autosomal recessive disorder with an occurrence rate of 1 in 12,000 to 40,000 live births. Common variable immunodeficiency (CVID) is an immune disorder that consists of hypogammaglobulinemia and impaired antibody response; although a TACI (transmembrane activator and calcium modulator and cyclophilin ligand interactor) receptor genetic mutation accounts for 10% of cases, the underlying defect among the rest is yet to be identified and may be multifactorial. Patients may present at any age, but the average reported age at onset of symptoms is 23 and 28 years for men and women, respectively. Despite its name, CVID occurs in only 1 in 25,000 to 60,000 persons, but it has been shown to be inherited as an autosomal recessive condition in some families. To our knowledge, there has never been a comorbidity consisting of PCD and CVID described in the literature.

CASE REPORT

A 17-year-old white boy was seen in consultation after hospitalization for the third time in 2 months with fever, malaise, chest tightness, emesis, and diarrhea. An offspring of nonconsanguineous parents delivered after a 32-week-gestation pregnancy, he was diagnosed with situs inversus at birth. After treatment for recurrent episodes of sinusitis and otitis media during the first 2 years of life, he underwent nasal biopsy, which showed abnormal ciliary structure on electron microscopy, consistent with a clinical picture of KS and PCD. Ultimately, the patient underwent myringotomy tube placement 6 times, adenoidectomy twice, and nasal septoplasty with polyp removal. Recurrent pneumonias were treated on an outpatient basis, and chest computed tomography (CT) at age 13 showed no evidence of bronchiectasis. The patient was up-to-date with respect to immunizations and experienced no adverse effects when immunized with live-virus vaccines. His growth patterns were reportedly normal. His family history was negative for immunodeficiency, and he denied tobacco use. His first hospitalization ever at age 15 was for pneumonia with pleural effusion, and his sputum culture grew ceftriax-
one-resistant *Streptococcus pneumoniae*, and repeat chest CT revealed left lower-lobe bronchiectasis.

Two months before this admission, while hospitalized for fever, emesis, diarrhea, and dehydration, the patient was noted to have significant splenomegaly accompanied by mild pancytopenia. Results of Epstein-Barr virus, cytomegalovirus, and toxoplasma studies were negative, and he was referred for a pediatric hematology/oncology evaluation of the pancytopenia and splenomegaly. He was readmitted to the hospital 1 week later with fever, cough, dyspnea, and pallor, and his chest CT now demonstrated significant bilateral interstitial consolidation, prominent mediastinal lymph nodes, and bilateral bronchiectasis. Bronchoscopy revealed copious tracheal secretions, and results of a bronchoalveolar lavage culture were positive for nonmucoid *Pseudomonas aeruginosa*. Bone marrow analysis was normal, and the patient’s pancytopenia was felt to be secondary to hypersplenism, which was present as a result of chronic infection.

On examination for the most recent admission, the patient appeared fatigued, pale, and chronically ill, with bilateral rales auscultated on pulmonary examination and the spleen tip palpable 5 cm below the right costal margin. A chest radiograph revealed left lower-lobe consolidation and right lower-lobe discoid atelectasis. The blood culture grew *S pneumoniae*, a complete blood count showed lymphopenia (absolute lymphocyte count: 1000 cells per μL), mild anemia (hemoglobin: 11.8 g/dL), and mild thrombocytopeny (120 000 platelets per μL), and quantitative immunoglobulin (Ig) values were IgA < 10 mg/dL, IgG < 50 mg/dL, and IgM < 10 mg/dL, with similar values on repeat determination. Subsequent outpatient workup showed nonprotective antibody titers to diphtheria, tetanus, and 12 pneumococcal serotypes, and only antidiptheria antibodies showed minimal response after booster immunizations with tetanus and diphtheria and 23-valent pneumococcal vaccines. A lymphocyte-enumeration study reflected the overall lymphopenia, but relative numbers of CD3, CD4, CD8, CD19, and CD16/56 cells were of normal proportion, and results of lymphocyte-proliferation studies to 3 mitogens were normal; HIV serology was negative, complement levels were normal, and stool giardia antigen was negative. Monthly intravenous immunoglobulin (IVIG) therapy was initiated.

**DISCUSSION**

Absent dynein arm structure leading to immotile cilia is the most common defect underlying the clinical features of KS, although other ciliary structural and functional defects have been identified. Absent ciliary function during embryogenesis is believed to result in situs inversus in 50% of cases, and male patients produce immotile sperm. Poor mucociliary clearance is deemed to be the primary reason behind recurrent sinopulmonary infections in KS and the broader classification of PCD, and the average age at diagnosis of PCD is 4.4 years. Although patients with PCD have been thought to be intact immunologically, a case report of an individual with KS and a neutrophil chemotactic defect, however, prompted Neffen et al to measure humoral and cellular immunity in a series of 6 patients with KS. In their study, immunoglobulin levels, complement assays, lymphocyte enumeration, and in vitro lymphoproliferative responses were essentially normal; however, an absent in vivo delayed hypersensitivity response to purified protein derivative and fungal antigens was observed in the majority of the patients. No other description of immunodeficiencies in patients with KS or PCD was identified during literature searches.

It has been found that recurrent pulmonary infections of several etiologies can lead to bronchiectasis. A study of 136 children with non–cystic fibrosis–related (non-CF) bronchiectasis reported PCD as the third most common etiology (15%), behind idiopathic (26%) and chronic (18%) aspiration. Although those with immunodeficiencies of all types comprised 35% of the study population, CVID was the most common individual immunodeficiency and was the fourth most common etiology (10%) of non-CF bronchiectasis overall. Other studies of non-CF bronchiectasis have shown similar statistics for PCD, although the prevalence of CVID ranged from 2% to 4%. Because of substantial overlap in the patterns of bronchiectasis seen in PCD and CVID, it is not possible to determine the etiology of bronchiectasis by chest CT findings alone. It has been noted, however, that pulmonary function in patients with PCD remains fairly constant if the condition is treated properly. A review of sequential pulmonary-function studies showed that our patient began to exhibit a decline after several years of stability. Meanwhile, a study of 248 patients with CVID described a 21% prevalence of gastrointestinal disease with a variety of pathologic findings on biopsy. Symptoms at 2 of our patient’s hospitalizations included emesis, diarrhea, and abdominal pain on presentation, and the persistence of these symptoms eventually led to the performance of upper and lower endoscopy. His biopsy specimens showed active esophagitis and mild active duodenitis. Although it has not been widely used clinically, analysis of exhaled nitric oxide and nasal nitric oxide showed significantly lower levels of both in patients with PCD compared with healthy controls, whereas subjects with non-CF bronchiectasis had higher levels than controls. Such testing was not available for our patient.

**CONCLUSIONS**

This case illustrates the need to consider associated diagnoses in complicated clinical situations. Although there was no available record of any humoral immunity evaluation having been performed during our patient’s early childhood, results of such testing more than likely
would have been normal given the fact that the usual age of onset of CVID is in the third decade of life. Once the relatively rare diagnosis of PCD was established, however, the possibility of an associated, equally rare immunodeficiency was considered only after a few years of clinical deterioration had occurred. Although the management of PCD includes early and aggressive use of antibiotics and immunization, the diagnosis of CVID is approached with antibiotics and IVIG replacement therapy. Immunizations are ordinarily of little benefit to patients with CVID, and live vaccines, in particular, can be harmful (and should be avoided). It is not possible to know the time of onset of CVID or whether immunoglobulin replacement earlier in this patient’s life might have delayed the progression of the structural lung disease; however, his clinical course improved after monthly IVIG administration. As this case illustrates, when faced with a deteriorating clinical picture, the clinician must consider comorbidities that require different or additional management strategies.

REFERENCES
Fever can precipitate ventricular tachycardia in adults with Brugada syndrome, but such a link has not been reported in children. A 21-month-old white girl presented repeatedly with decreased conscious level and seizures during fever. During a typical episode, rapid ventricular tachycardia was documented. The resting 12-lead electrocardiogram revealed a Brugada electrocardiogram signature. Resting electrocardiograms of the asymptomatic brother and mother were normal, but fever in the mother and pharmacologic stress with ajmaline in the brother revealed Brugada electrocardiogram features. Genetic testing revealed an SCN5A mutation in the affected family members.

**CASE REPORT**

A 21-month-old previously healthy white girl presented to the emergency department during the night after a third episode of seizure with a high fever. The patient had been seen at the same hospital 3 months earlier (during the second attack) with a high fever (38.5°C) that was diagnosed as a febrile seizure. On all 3 occasions, occurring some months apart, she woke the parents with a cry of distress; they found her in her cot with the sheet covering her while she was jerking rhythmically. She was “either asleep or unconscious.” She felt...
very hot, and after uncovering her, the jerking, which affected both arms, rapidly ceased. After this she was limp, very pale, and unresponsive, but she was breathing. With efforts to cool her, her conscious level increased over a few minutes. During a third episode, described by the parents as identical to the previous 2 but lasting longer despite cooling, she presented to the emergency department.

She was pale and unwell; her temperature was 38.0°C and pulse rate was 200 beats per minute, with a gallop rhythm and poor peripheral perfusion. She was limp, distressed, and whimpering, not vocalizing clearly, and did not recognize her parents as being distinct from nursing staff.

There were clinical signs of an upper respiratory tract infection. The precordium was noted to be hyperactive. The 12-lead ECG showed a rapid, broad-complex tachycardia (Fig 1). The patient received 3 bolus doses of adenosine (50, 100, and 250 μg/kg), to no effect. After a cough, her sinus rhythm returned (Fig 2A). Immediately her color improved and she sat up, talked to her parents, and started naming objects in the room. She was started on oral sotalol 2 mg/kg twice daily. The fever took some days to settle and was managed with paracetomol; no additional seizures or tachycardias occurred. Brugada syndrome was suspected from the resting ECG (Fig 2A), and she was transferred to a pediatric cardiology unit for additional investigation.

The working diagnosis, therefore, was that she had had reduced conscious level and seizure triggered by low cerebral perfusion as a result of a fever-triggered ventricular tachycardia.

The family history revealed no incidences of seizure, sudden death, or syncope, and resting ECGs on the brother and both parents were normal. An echocardiogram and cardiac MRI scan were normal. An intravenous challenge with ajmaline, a short-acting sodium-channel–blocking agent, was performed. The result was strongly positive, with very marked exaggeration of the precordial ST segment elevation at low doses (0.2 mg/kg) (Fig 2B). The family was provided with an automatic external defibrillator for home use, given open access to the local hospital, and advised to treat fevers aggressively with paracetomol and tepid sponging. The sotalol was continued because of its apparent early therapeutic benefit, with no recurrence of seizures or ventricular tachycardia with subsequent fevers over the first few days. Before planned ajmaline tests on the parents, both of whom had normal resting ECGs, the mother had some palpitations during an attack of influenza. At the time of high fever, her ECG revealed typical Brugada features (Fig 3). An invasive electrophysiology study did not induce ventricular arrhythmias. The brother also had a normal resting ECG, but his ajmaline test was also positive, although the ECG changes occurred at higher doses than those of his sister (0.8 mg/kg).

GENETIC ANALYSIS OF SCN5A
After obtaining informed consent for genetic testing, blood samples were taken from the presenting child for

![FIGURE 1](image)
mutation analysis of SCN5A. The genomic DNA was used for polymerase chain reaction–based assays covering the coding and flanking intronic regions of the gene, and the polymerase chain reaction products were screened for heterozygous profiles by using conventional denaturing high-pressure liquid chromatography (dHPLC) technology (Transgenomic, Omaha, NE). The output profiles from the dHPLC were scrutinized for abnormal profiles indicating the presence of heterogeneous DNA mismatching. Abnormal dHPLC profiles were sequenced in both directions (Applied Biosystems, Foster City, CA), and sequence changes were verified by restriction fragment length polymorphism (RFLP) analysis when available. RFLP studies of mutation frequencies were conducted on 2% molecular screening agarose gels (Roche, Indianapolis, IN) using 150 random normal-control cases to evaluate candidate pathogenic mutations within the general population.

RESULTS

Genetic Testing
The dHPLC screening revealed several abnormal profiles that required sequencing. Sequencing of SCN5A exon 24 demonstrated a novel, heterozygous 1-base pair (bp) insertion mutation (InsG) in position nt4392–4396 of the SCN5A coding sequence (Fig 4). This mutation was also found in the mother and brother but not in the father. In the absence of differentiating restriction-enzyme–digest profiles, the mutation was excluded from
300 normal-control chromosomes by dHPLC profiling (not shown). No other mutations were found.

Clinical Outcome

The girl was continued on sotalol, and an advisory external defibrillator was provided for the family. Fevers were managed with paracetamol. Three and a half years later, the girl remains well, having had many febrile episodes and no more seizures or clinically evident tachycardias. Her resting ECG remains abnormal. No intervention was given for the mother or brother, who both remain well. Dislocation of the genetic maternal family has thwarted efforts to screen more distant family relatives for Brugada syndrome.

DISCUSSION

Brugada syndrome most commonly causes sudden unexpected death in young Asian men during sleep. It is inherited in an autosomal dominant manner with variable clinical expression, and females are significantly less likely to die suddenly. We have not been able to find a single case report of death after a febrile seizure, which suggests that Brugada syndrome, with a fever-triggered ventricular arrhythmia causing decreased conscious level during fever, must be a very rare mimic of this usually benign condition. Nevertheless, ECGs are not typically a routine part of clinical assessment after a seizure associated with fever, and the condition may be more frequent than suspected. Furthermore, there may be a reluctance by forensic pathologists to attribute deaths during or after seizures to febrile seizure because of their benign nature by definition. Deaths with seizures secondary to long QT syndrome have been erroneously ascribed to epilepsy in the past.

The Brugada syndrome mutations in SCN5A cause a dominant negative effect or loss of function in the sodium current, which is a determinant of the phase 0 and phase 1 segments of the cardiac action potential. Carriers of Brugada syndrome can have a normal resting ECG, but recent observations in adults have demonstrated that classical Brugada ECG features, even ventricular tachycardia and death, can be precipitated by fever. No such link to fever has been shown in children, yet 3 independent cases of infants with ventricular fibrillation or ventricular tachycardia have been reported recently in association with mutations in SCN5A: 2 associated with long QT syndrome type 3 and one in association with Brugada syndrome.

In our patient, the 1-bp insertion causes the reading frame of SCN5A to be disrupted; a premature stop codon (at position 1483 of SCN5A[1483X]) results in a truncated SCN5A polypeptide. Other such SCN5A mutations that generate a premature termination codon are asso-
associated with distinct clinical phenotypes; 4196delA
(V1397X) is associated with Brugada syndrome, and
5280delG(1768X) is associated with conduction dis-
ease.24,25 Electrophysiological studies of both of these
mutants failed to express any sodium currents; thus, it is
likely that the SCN5A mutation in the present case will
also fail to produce sodium currents. Although cellular
electrophysiology has not been performed, the ECG
changes in the family are so characteristic that the Bru-
gada phenotype is not in doubt.

This case is the first reported incident in which the
febrile-onset phenomenon has occurred in a young child
and is made more remarkable by the fact that she is both
female and white. Overheating is a known risk factor for
sudden infant death syndrome and the SCN5A gene is
the most common of any to be linked to sudden infant
death.26 It is tempting to speculate that this correlation
may be attributable in part to heat-triggered ventricular
tachycardia or fibrillation in such genetically vulnerable
infants.

Our patient and her mother have a fever-dependent
expression of the phenotype. In the case of our patient,
the ventricular tachycardia may have happened during a
critical developmental window, as reflected by the ab-
sence of ventricular tachycardia or seizures during sub-
sequent febrile illness. A family with young-age–specific
sudden death was described recently with a Brugada
SCN5A mutation.27 It is also possible that sotalol therapy
or liberal use of antipyretics have stopped a recurrence.
No medications (including sotalol), with the emerging
possible exception of quinidine,28 have proven therapeu-
tic benefit for Brugada syndrome. Although pure β
blockers, in large doses, can unmask Brugada ECG
changes and may be arrhythmogenic,29 we have been
reluctant to discontinue sotalol (which has both
β-blocking and repolarization-prolonging [Vaughan-Williams
type III] properties) because of the apparent beneficial
effect on arrhythmia prevention during our patient’s
many subsequent high fevers over >3 years.

Molecular defects in Na+1 channels have been re-
ported in association with temperature-sensitive disor-
ders such as generalized epilepsy with febrile seizures
plus and heat-induced myotonia and cold-induced pa-
ralysis in congenital paramyotonia.30–32 Furthermore, a
patient with a febrile illness and cardiac conduction dis-
ease who experienced a syncope was reported to be a
carrier for a SCN5A mutation (F1344S); cellular electrophysiological testing showed a tempera-
ture-sensitive sodium current.33 SCN9A (Na+,1.7) muta-
tions have also been reported in primary erythromyi-
a, a rare disorder that is characterized by intermittent
burning pain with redness and heat in the extremi-
ties.35,36

CONCLUSIONS
Fever can trigger ventricular tachycardia in young chil-
dren with previously unrecognized Brugada syndrome,
and the low-output state can trigger a seizure, the pre-
sentation thus mimicking febrile seizures.

This case serves as a reminder that occult cardiac
channelopathies should be borne in mind in the inves-
tigation of childhood seizures with or without fever.
However, we would not yet advise that all children who
have a seizure or reduced conscious level during fever
should have an ECG. It would be prudent to order an
ECG if there are clinical features suggesting a cardiac
arrhythmia (tachycardia out of proportion to the fever,
weak peripheral pulses, hyperactive precordium, and
marked pallor) or a family history suggesting Brugada
syndrome, such as young sudden death, particularly in
males at night.

ACKNOWLEDGMENTS
The molecular genetics work was enabled by grants from
Cure Kids, the Lion Foundation, the Greenlane Research
and Education Fund, and the John Neutze Fund.

We thank Dr Iain Melton for providing information
and ECGs on the parents, and we thank the parents of
the child concerned, who enthusiastically supported the
publication of this report.

REFERENCES
Child. 2004;89:751–756
2. Nelson KB, Ellenberg JH. Prognosis in children with febrile
3. Verity CM, Golding J. Risk of epilepsy after febrile convulsions:
a national cohort study [published correction appears in BMJ.
4. Gordon N. Sudden unexpected death in epilepsy. Dev Med Child
Neurol. 2001;43:354–357
5. Priori SG, Napolitano C. Genetics of cardiac arrhythmias and
6. Dumaine R, Antzelevitch C. Molecular mechanisms underlying
the long QT syndrome. Curr Opin Cardiol. 2002;17:36–42
a mechanism of acceleration of ventricular tachycardia. Cir-
culation. 1990;81:1633–1643
8. Wang Q, Shen J, Splawski I, et al. SCN5A mutations associated
with an inherited cardiac arrhythmia, long QT syndrome. Cell.
1995;80:805–811
Brugada syndrome. Eur Heart J. 2002;23:675–676
10. Napolitano C, Rivolta I, Priori SG. Cardiac sodium channel
ical basis of sudden unexplained nocturnal death syndrome
(SUNDS), a disease allelic to Brugada syndrome. Hum Mol
Genet. 2002;11:337–345
challenge using flecainide and ajmaline in patients with Bru-
marker for the sudden unexplained death syndrome in


ABSTRACT

A rare complication of infection with the Epstein-Barr virus is the development of hemophagocytic lymphohistiocytosis. Although most cases of Epstein-Barr virus–induced hemophagocytic lymphohistiocytosis develop in immunocompetent individuals, the rare immunodeficiency X-linked lymphoproliferative disease is often unmasked by Epstein-Barr virus infection and is clinically indistinguishable from Epstein-Barr virus–induced hemophagocytic lymphohistiocytosis. We describe the clinical course and management of a previously healthy 17-year-old boy who presented with hemodynamic collapse and severe systemic inflammatory response syndrome resulting from overwhelming hemophagocytosis in the setting of X-linked lymphoproliferative disease. A novel therapeutic approach using anti–tumor necrosis factor α (TNF-α) therapy was instituted, aimed at attenuating the viral-induced hyperinflammatory state. Given the similarity to overwhelming sepsis, yet a substantially different therapeutic approach, this case illustrates the importance of early recognition and prompt treatment that are necessary to reduce the high morbidity and mortality associated with Epstein-Barr virus–induced hemophagocytic lymphohistiocytosis and X-linked lymphoproliferative disease.

EPSTEIN-BARR VIRUS (EBV), which preferentially infects B cells, has a variety of clinical presentations that range from an asymptomatic carrier state to a fatal overwhelming infection. In the majority of the population, EBV causes acute infectious mononucleosis, a self-limiting illness characterized by fever, lymphadenopathy, tonsilopharyngitis, and hepatosplenomegaly. An infrequent complication of EBV infection is the development of hemophagocytic lymphohistiocytosis (HLH), a disorder of unregulated activation of lymphohistiocytic cells that results in hemophagocytosis, hypercytokinemia, and multiorgan system dysfunction. Although associated with substantial morbidity and mortality, early recognition and prompt therapy may result in successful treatment of EBV-induced HLH.

Here we present the case of a previously healthy 17-year-old boy with acute infectious mononucleosis followed by a rapidly deteriorating clinical course resulting from EBV-induced HLH and associated with underlying immunodeficiency X-linked lymphoproliferative (XLP) disease. We also describe our therapeutic approach, which incorporated a novel application of anti–tumor necrosis factor α (TNF-α) therapy.

CASE REPORT

Four weeks before admission, a previously healthy 17-year-old boy was diagnosed with infectious mononucleosis. He presented with fever, lymphadenopathy, tonsillitis, and hepatosplenomegaly. Laboratory findings revealed marked leukopenia, thrombocytopenia, and elevated liver enzymes. Despite aggressive supportive care, his condition rapidly deteriorated, with development of hemodynamic instability and multiorgan failure.

Key Words: X-linked lymphoproliferative disease, hemophagocytic lymphohistiocytosis, systemic inflammatory response syndrome, infectious mononucleosis, sepsis, Epstein-Barr virus, Etanercept

Abbreviations: EBV, Epstein-Barr virus; HLH, hemophagocytic lymphohistiocytosis; XLP, X-linked lymphoproliferative; TNF-α, tumor necrosis factor α; IL, interleukin; SAP, signaling lymphocyte activation molecule–associated protein; Th, T helper; LMP-1, latent membrane protein 1; IVIG, intravenous immunoglobulin

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1534
doi:10.1542/peds.2006-1534

Accepted for publication Oct 31, 2006
Address correspondence to Timothy T. Cornell, MD, Division of Pediatric Critical Care Medicine, C.S. Mott Children’s Hospital, University of Michigan, 1580 E Medical Center, Ann Arbor, MI 48109. E-mail: ttcornel@med.umich.edu

PEDIATRICS (ISSN Numbers: Print, 0031–4005; Online, 1098–4275). Copyright © 2007 by the American Academy of Pediatrics
osis on the basis of a 1-week history of fever, malaise, myalgias, headache, abdominal pain, tender splenomegaly, and emesis. EBV antibody titers showed acute infection, with a complete blood cell count showing mild pancytopenia and atypical lymphocytes consistent with acute infectious mononucleosis. In the weeks after diagnosis his symptoms persisted, with recurrent fever to 39°C, nausea, abdominal pain, pharyngitis, intermittent truncal maculopapular rash, fatigue, and poor oral intake. Four weeks after his initial diagnosis, he was readmitted to the hospital for worsening pancytopenia and dehydration. Despite transient response to aggressive intravenous rehydration, his clinical condition rapidly deteriorated on hospital day 2, with refractory hypotension, decreasing urine output, and declining mental status. He was subsequently transferred to the PICU.

On arrival to our PICU, his examination was consistent with low cardiac output including poor distal perfusion and purplish discoloration noted at the distal tips of his fingers, toes, and nose. His admission laboratory values are listed in Table 1. He subsequently went into cardiorespiratory failure and required intubation and mechanical ventilation along with inotropic and vasopressor therapy. Given his clinical presentation of fever, hypotension, and pancytopenia, broad-spectrum antibiotics were empirically started for concern of severe sepsis. Because previous viral infections are commonly associated with subsequent secondary bacterial infections (often termed “superinfections”), this distinct possibility was considered as a principal cause of his acute cardiorespiratory collapse. Blood and urine cultures were performed on presentation and throughout his hospitalization, and results remained negative for the duration of his stay.

His prolonged course of EBV, however, led to the inclusion of EBV-related complications such as HLH in the differential diagnosis for the patient’s rapid decline. A bone marrow biopsy was performed soon after admission to the PICU, with findings concerning for EBV-induced HLH, namely, pancytopenia, widespread hemophagocytosis, and large granular lymphocytes. Concurrent elevations in lactate dehydrogenase, ferritin, EBV-DNA copy numbers, and soluble interleukin (IL) 2 receptor α (Table 1) provided additional supportive evidence for EBV-induced HLH. Genetic studies performed at one of the author’s (Dr Filipovich’s) HLH reference laboratory showed normal perforin activity consistent with secondary HLH. Additional testing showed decreased natural killer cell function and absent signaling lymphocyte activation molecule–associated protein (SAP), both of which are consistent with the underlying immunodeficiency XLP disease. Subsequent genetic testing of the SH2D1A gene, which encodes for SAP, revealed no mutations. Furthermore, a detailed family history revealed no XLP disease or immunodeficiency in either the maternal or paternal lineage, which suggested the diagnosis of “sporadic” XLP disease with the specific genetic defect yet undefined versus sporadic HLH with absent SAP and natural killer cell function.

His initial treatment regimen consisted of intravenous immunoglobulin, high-dose methylprednisolone, and rituximab, an anti-CD20 antibody. However, over the next 36 hours, his systemic inflammatory-response syndrome worsened and led to multiple organ dysfunction, as reflected by the need for continuous venovenous hemofiltration therapy for renal failure, increased blood-product replacement for hematologic failure, massive inotropic and vasopressor support for cardiovascular failure, and continued mechanical ventilatory support for respiratory failure.

Given the uncontrolled immune activation unrelated to bacterial infection, etoposide therapy was instituted at 150 mg/m² in accordance with the HLH-2004 treatment protocol. His clinical status continued to show little improvement, and the decision was made to administer an anti–TNF-α antibody: etanercept 25 mg intravenously. After his dose of etanercept, his clinical status stabilized and improved swiftly over the next 36 hours, with improved liver and kidney function, a decrease in

---

**TABLE 1** Laboratory Values on Admission to the PICU

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>White blood cell count</td>
<td>1600/mm³</td>
</tr>
<tr>
<td>Neutrophils</td>
<td>84.3%</td>
</tr>
<tr>
<td>Lymphocytes</td>
<td>10.5%</td>
</tr>
<tr>
<td>Monocytes</td>
<td>5.1%</td>
</tr>
<tr>
<td>Eosinophils</td>
<td>0.1%</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>13.1 g/dL</td>
</tr>
<tr>
<td>Hematocrit</td>
<td>36.9%</td>
</tr>
<tr>
<td>Platelet count</td>
<td>20 000/mm³</td>
</tr>
<tr>
<td>Lactate dehydrogenase</td>
<td>988 IU/L</td>
</tr>
<tr>
<td>Sodium</td>
<td>127 mg/dL</td>
</tr>
<tr>
<td>Potassium</td>
<td>4.8 mg/dL</td>
</tr>
<tr>
<td>Chloride</td>
<td>101 mg/dL</td>
</tr>
<tr>
<td>Bicarbonate</td>
<td>17 mg/dL</td>
</tr>
<tr>
<td>Urea nitrogen</td>
<td>22 mg/dL</td>
</tr>
<tr>
<td>Creatinine</td>
<td>1.9 mg/dL</td>
</tr>
<tr>
<td>Glucose</td>
<td>88 mg/dL</td>
</tr>
<tr>
<td>Calcium</td>
<td>6.9 mg/dL</td>
</tr>
<tr>
<td>Protein</td>
<td>5.2 g/dL</td>
</tr>
<tr>
<td>Albumin</td>
<td>2.2 g/dL</td>
</tr>
<tr>
<td>Aspartate amino transferase</td>
<td>179 IU/L</td>
</tr>
<tr>
<td>Alanine amino transferase</td>
<td>100 IU/L</td>
</tr>
<tr>
<td>Alkaline phosphatase</td>
<td>266 IU/L</td>
</tr>
<tr>
<td>Bilirubin, total</td>
<td>6.5 mg/dL</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>89 mg/dL</td>
</tr>
<tr>
<td>International normalized ratio</td>
<td>2.0</td>
</tr>
<tr>
<td>Ferritin</td>
<td>7360 ng/mL</td>
</tr>
<tr>
<td>D-dimer</td>
<td>18.7 mg/L</td>
</tr>
<tr>
<td>Soluble IL-2 receptor</td>
<td>85523 pg/mL</td>
</tr>
<tr>
<td>EBV-DNA peripheral whole blood</td>
<td>26000 copies per mL</td>
</tr>
<tr>
<td>Immunoglobulins (before IVIG)</td>
<td></td>
</tr>
<tr>
<td>IgG</td>
<td>797 mg/dL</td>
</tr>
<tr>
<td>IgA</td>
<td>84 mg/dL</td>
</tr>
<tr>
<td>IgM</td>
<td>120 mg/dL</td>
</tr>
</tbody>
</table>

Values were obtained from a bone marrow biopsy: marrow cellularity, ~80%, with granulocytic and megakaryocytic hyperplasia, near aplasia of erythroid series, marked interstitial increase in histiocytes, and hemophagocytic histiocytes.

---

PEDIATRICS Volume 119, Number 5, May 2007 e1213
blood-product replacement, and a decrease in inotropic and vasopressor support (Fig 1). On the basis of the favorable clinical response observed after the initial dose but in light of ongoing organ dysfunction, the decision was made by the clinical care team to administer a subsequent course of etanercept. The same dose was administered and was associated with resolution of vasopressor and inotropic requirement, decrease in mechanical ventilatory requirements, and improvement in renal function, which led to eventual withdrawal of continuous renal replacement therapy.

Levels of ferritin and lactate dehydrogenase, which have been suggested to be reliable clinical markers of disease activity in HLH, gradually decreased over the subsequent week, with repeat EBV DNA showing 0 copies on intensive care day 7 correlating with his white blood cell count. His fever curve also began to improve and gradually normalized by day 14 of intensive care.

In the 2 weeks after starting etanercept and etoposide, the patient demonstrated steady resolution of his multiple organ failure, affording the discontinuation of renal replacement therapy and continuous infusions of vasoactive and sedative agents. Despite discontinuation of his sedation, the patient showed minimal signs of cognitive function. At this time, initial brain imaging was performed using MRI, which revealed a nonspecific inflammatory pattern with scattered, nonenhancing signal abnormalities in the frontal white matter and cerebellar cortex. The hematologic and clinical abnormalities often suggest sepsis, leukemia, lymphoma, or systemic autoimmune vasculitis as the underlying cause. The true prevalence of the disease is unknown at this time, because it remains largely underdiagnosed in Western countries. Although the majority of EBV-induced HLH develops in immunocompetent individuals, EBV can unmask an underlying immune disorder, the most common being XLP disease. The phenotypic presentation of EBV-induced HLH and XLP disease is largely indistinguishable, and both carry a reported mortality rate of >90%.

XLP disease is a congenital immunodeficiency that is estimated to affect 1 in 1 000 000 males that may

DISCUSSION
HLH was first described in 1939 and has since been broken down into primary and secondary HLH. Primary HLH, also termed familial HLH, is an autosomal recessive disease with an identical phenotype to secondary HLH. More than 70% of patients with familial HLH develop the disease at <1 year of age, although familial forms have been reported into early adulthood. It is linked with mutations in the gene coding for perforin, which was evaluated and found to be normal in our patient. Secondary HLH has been associated with immunologic triggers including malignancies, bacterial or parasitic infections, and, most commonly, viral infections including EBV, cytomegalovirus, parvovirus, and HIV. The prototype and most often reported association is EBV-induced HLH. EBV-induced HLH can affect any age group, ranging from infants to young adults, and tends to occur in apparently immunocompetent individuals.

EBV-induced HLH is a distinct clinical entity with defined diagnostic criteria (Table 2) characterized by evidence of EBV infection, persistent fever, cytopenia, liver dysfunction, coagulopathy, hepatosplenomegaly, and hemophagocytosis in the bone marrow, lymph nodes, liver, and spleen. As the signs and symptoms of EBV-induced HLH imply, the differential diagnosis is extremely broad, particularly in the ICU setting with a critically ill patient. The hematologic and clinical abnormalities often suggest sepsis, leukemia, lymphoma, or systemic autoimmune vasculitis as the underlying cause. The true prevalence of the disease is unknown at this time, because it remains largely underdiagnosed in Western countries. Although the majority of EBV-induced HLH develops in immunocompetent individuals, EBV can unmask an underlying immune disorder, the most common being XLP disease. The phenotypic presentation of EBV-induced HLH and XLP disease is largely indistinguishable, and both carry a reported mortality rate of >90%.

XLP disease is a congenital immunodeficiency that is estimated to affect 1 in 1 000 000 males that may

![FIGURE 1](image)

**FIGURE 1**
Timing of etanercept dose and vasopressor/inotropic requirements.

### TABLE 2: HLH Diagnostic Criteria

<table>
<thead>
<tr>
<th>Major criteria</th>
<th>Alternative criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fever of &gt;38° C for at least 7 d</td>
<td>Low or absent natural killer cell activity</td>
</tr>
<tr>
<td>Splenomegaly</td>
<td>Serum ferritin level of &gt;500 µg/L</td>
</tr>
<tr>
<td>Cytopenia involving ≥2 cell lines</td>
<td>Soluble CD25 (soluble IL-2 receptor) level at &gt;2400 U/mL</td>
</tr>
<tr>
<td>Hypertriglyceridemia or hypofibrinogenemia</td>
<td></td>
</tr>
<tr>
<td>Hemophagocytosis demonstrated in bone marrow, spleen, or lymph nodes without evidence of malignancy</td>
<td></td>
</tr>
</tbody>
</table>

*a* A diagnosis of HLH requires the presence of all 5 major criteria. If the patient meets only 4 criteria but the clinical suspicion for HLH is high, one should initiate treatment, because delays may be fatal.

*b* Alternative criteria 1 or a combination of 2 and 3 may substitute for 1 major criterion.

present anywhere from childhood to early adulthood. The 3 main phenotypes of XLP disease are (1) inappropriate immune response to EBV with fatal or near-fatal infectious mononucleosis, (2) lymphoproliferative disorders, typically of B-cell origin, and (3) dysgammaglobulinemia. The most common presenting phenotype is fulminant infectious mononucleosis, which is highly fatal and clinically and pathophysiologically identical to EBV-induced HLH. Indeed, the diagnosis of EBV-induced HLH should prompt an investigation for XLP disease, because the clinical presentation of EBV-induced HLH is the most common manifestation of underlying XLP disease.6–9 In addition to fulminant infectious mononucleosis, our patient also exhibited dysgammaglobulinemia after presentation to the PICU, which is consistent with the underlying diagnosis.

The genetic defect in XLP disease was recently mapped to Xq25 in the SH2D1A gene. The SH2D1A gene is responsible for an SAP, found to be deficient or absent in patients with XLP disease.5,13,14 The SH2 domain of SAP binds a phosphorylated tyrosine in the cytoplasmic domain of 2 type 1 transmembrane receptors, physically associating with signaling lymphocyte activation molecule (SLAM) receptors present on T and B lymphocytes and 2B4 receptors present on natural killer cells.15–17 There is some evidence that SAP-deficient cytotoxic T cells are unable to lyse EBV-positive B cells, leading to the rapid proliferation of large granular lymphocytes and enhanced T helper 1 (Th1) cytokine production, which leads to the cytokine storm seen in EBV-induced HLH and XLP disease.18 Other studies suggest deficiency in natural killer cell lytic activity or lack of development of a subset of cytotoxic T lymphocytes known as natural killer T cells.19–21 Regardless of the role of SAP, genetic sequencing of the SH2D1A gene is available to detect known mutations that produce XLP-disease phenotypes and should be sent whenever fulminant infectious mononucleosis and EBV-induced HLH is encountered. The gene testing can detect 97% of known SH2D1A mutations in affected males with the XLP-disease phenotype, but only in those with ≥2 maternal relatives with XLP disease.6,9 Our patient had no family history of XLP disease, and his SH2D1A gene testing revealed none of the known mutations in the coding sequence. Testing of his SAP activity on flow cytometry, however, revealed nearly absent activity and, coupled with his clinical presentation of fulminant infectious mononucleosis and dysgammaglobulinemia, led to the diagnosis of XLP disease rather than sporadic secondary EBV-induced HLH.

The exact mechanism of how EBV induces HLH, with or without XLP disease and producing the characteristic clinical response, has only recently begun to be elucidated. It has been suggested that EBV infection of B cells triggers a polyclonal proliferation of cytotoxic T lymphocytes, which in turn stimulate histiocytes and macrophages, resulting in uncontrolled immune activation and subsequent hypercytokinemia.22,23 Research performed in chronic active EBV infection, a chronic ongoing mononucleosis of unknown etiology but with a somewhat similar presentation as that of EBV-induced HLH and XLP disease, has shown that activated T lymphocytes laden with EBV DNA express both Th1 and Th2 cytokines, which leads to ongoing inflammation and disease activity.1 Other data suggest that EBV targets CD8+ T cells and natural killer cells, which leads to rapid, uncontrolled proliferation and profound release of immense levels of IL-2, interferon γ, TNF-α, and IL-6, among other inflammatory cytokines, as a result of widespread lymphohistiocytic activation.24–27 More recent studies have focused on characteristics of the virus itself that induce the phenotype of XLP disease and HLH, one of which is latent membrane protein 1 (LMP-1) on the surface of EBV. In EBV-infected cell lines, LMP-1 was shown to upregulate the TNF-α gene and lead to increased secretion from infected T lymphocytes, corresponding with previous data showing elevations in TNF-α and macrophage activation in EBV-associated T-cell lymphoproliferative disorders.28,29 There is also evidence that LMP-1 may directly inhibit expression of the SAP gene, which leads to loss of the gene product and increased Th1 cytokine activation. In addition to therapeutic implications, this function of LMP-1 would provide a basis for the similar phenotypic presentation of XLP disease and HLH via gene mutation and suppression of the gene product, respectively.10

Our treatment plan had 2 aspects aimed at reducing the uncontrolled immune response. First, we targeted the sources of immune activation with intravenous immunoglobulin (IVIG) and high-dose corticosteroids followed by the institution of the HLH-2004 treatment protocol. The HLH-2004 protocol, which evolved from the HLH-94 protocol, was devised with the intention of interrupting the inappropriate immune response and cytokine-activation cascade by using immunosuppression. After a diagnosis of HLH or a clinical scenario consistent with the diagnosis (see Table 2 for diagnostic criteria), initial therapy involves etoposide at 150 mg/m^2 twice weekly and dexamethasone infusion at 10 mg/m^2 twice weekly. In the evolution of the HLH protocols, HLH-2004 now includes cyclosporin A, aiming for trough levels of 200 μg/L during the first week, assuming normal renal function.10 The use of cyclosporin is thought to improve patient outcomes by controlling the deregulated release of cytokines and improving neutrophil recovery during disease or therapy-induced neutropenia.4,31 Our patient was in acute renal failure on presentation, precluding the use of cyclosporin until renal function had returned to baseline. Finally, the HLH-2004 protocol also includes intrathecal methotrexate and prednisolone during initiation for patients with central nervous system symptoms that are documented either by irregular lumbar puncture or findings on central
nervous system imaging. Our patient did indeed show signs of central nervous system involvement, with cognitive impairment noted on presentation to the PICU and continuing after other organ systems had recovered. He did not show full cognitive recovery with intact sensorium until 6 weeks after his initial presentation, corresponding with improvement in the signal abnormality seen on his MRI. During initial immunochemotherapy, disease progression can be assessed with recovery of platelet counts, decrease in ferritin and soluble IL-2 levels, and absence of hemophagocytosis in bone marrow. In addition, elimination of EBV DNA documented by quantitative polymerase chain reaction has been shown to be useful in documenting response to therapy and predicting mortality. Our patient’s titers from peripheral whole blood responded well as therapy progressed and corresponded with improvement in his clinical status (Fig 2). The use of IVIG and initial high-dose corticosteroid therapy is well documented and was the standard of care before the initiation of the HLH-2004 protocol. Later studies have shown that patients who receive etoposide within 4 weeks of diagnosis have a higher rate of survival than those who receive only IVIG and steroids. However, there is still thought to be a role for IVIG in the initial presentation of EBV-induced HLH and XLP disease; indeed, IVIG is required once every 4 weeks in the HLH-2004 protocol. For our patient, IVIG and high-dose corticosteroids were initiated in conjunction with the bone marrow biopsy, when the diagnosis was still not entirely clear but highly suspected, not in lieu of the HLH-2004 protocol. In addition, although not considered standard therapy, the use of Rituxan in our patient was supported by recent data suggesting that B cells may also be targeted in EBV-associated XLP disease, and directed therapy with the anti-CD20 antibody may reduce morbidity and mortality by reducing the circulating B-cell population and, thus, EBV viral load. Of note, this approach will not target EBV-infected T lymphocytes that are thought to contribute to widespread immune activation as reviewed above.

The interruption of the immune activation with the HLH-2004 protocol takes several days to weeks to occur, and because of the rapid clinical deterioration of our patient, our second therapy was aimed at reducing the proinflammatory cytokine levels. The upregulation of the TNF-α gene leading to widespread macrophage activation has been documented in EBV infection of T cells, which suggests a possible role for the use of an anti-TNF-α antibody such as etanercept in HLH and XLP disease. Although etanercept has been suggested as a possible therapeutic intervention, its successful use has not been documented in the literature to date. Our decision to use etanercept was also supported by previous experience with it at our institution for noninfectious cytokine storm as well as data suggesting reduction in early mortality with anti-TNF-α therapy in sepsis. After administration of etanercept, our patient had a dramatic improvement in hemodynamics as reflected by a rapid decrease in vasopressor requirements, as well as improvement in renal and liver function and decreased requirement for blood-product replacement. However, we did not measure serum TNF-α before and after the institution of anti-TNF-α therapy, which makes it impossible to conclude unequivocally that etanercept specifically afforded control of the severe systemic inflammatory response exclusive of the effects

FIGURE 2
Association between EBV-genome copies per mL whole blood and white blood cell (WBC) count. The patient received methylprednisolone and IVIG on admission to the PICU. He received etoposide on intensive care days 1, 5, and 23, rituximab on day 1, and etanercept on days 2 and 7.
of the additional therapies including etoposide, IVIG, steroids, Rituxan, and/or continuous renal replacement therapy. The clinical response observed after anti–TNF-α antibody infusion strongly suggests this to be the case, especially in light of the known role of TNF-α in HLH and XLP disease. Regrettably, TNF-α levels were not obtained before or after etanercept therapy, so we are unable to precisely show TNF-α neutralization in this case. Nevertheless, additional studies with etanercept are needed in patients with EBV-induced HLH and XLP disease and other illnesses that result in nonbacterial-induced hypercytokinemia before its routine use in such patients can be universally recommended.

CONCLUSIONS
Acute EBV infection is not life-threatening in the majority of individuals who are infected. Rarely, however, life-threatening complications develop that require immediate recognition and treatment. In the presence of prolonged EBV infection and severe hemodynamic collapse with pancytopenia, coagulopathy, and hepatosplenomegaly, EBV-induced HLH should be suspected and warrant determination of ferritin levels, EBV studies, and a bone marrow aspiration to expedite diagnosis and direct life-saving therapy. Furthermore, any case of EBV-induced HLH in males should prompt an investigation for XLP disease, because the two are clinically indistinguishable. The HLH-2004 protocol provides guidelines for management of these patients in the acute setting, and rapid therapy aimed at attenuating the proinflammatory mediators using etanercept may be warranted given previous research and our experience in this case. As targeted therapy evolves, increased awareness of EBV-induced HLH and XLP disease is necessary to improve and prevent the high associated morbidity and mortality.

REFERENCES
17. Shlapatska LM, Mishkalap SV, Berdova AG, et al. CD150 association with either the SH2-containing inositol phosphatase or the SH2-containing protein tyrosine phosphatase is regulated by the adaptor protein SH2D1A. J Immunol. 2001;166:5480–5487
25. Kasahara Y, Yachie A. Cell type specific infection of Epstein-Barr virus (EBV) in EBV-associated hemophagocytic lympho-
histiocytosis and chronic active EBV infection. Crit Rev Oncol Hematol. 2002;44:283–294